



UNIVERSITÉ FRANÇOIS - RABELAIS DE TOURS



ÉCOLE DOCTORALE Santé, Sciences, Technologies
ÉQUIPE Transcription, Lymphome Viro-Induit

THÈSE présentée par :

Damien COUPEAU

soutenue le : 09 juin 2011

pour obtenir le grade de : **Docteur de l'université François – Rabelais de Tours**

Discipline/ Spécialité : Sciences de la vie/Virologie

Etude transcriptionnelle et post-transcriptionnelle de la région IR_L/TR_L du virus de la maladie de Marek

THÈSE dirigée par :

M RASSCHAERT Denis

Professeur, Université François - Rabelais, Tours

RAPPORTEURS :

M PFEFFER Sébastien

Chargé de recherche/HDR, CNRS/Strasbourg

M LEGRAND Alain

Professeur, Université d'Orléans

JURY :

Mme DUPUY Catherine

Maître de conférence, Université François - Rabelais, Tours

M LEGRAND Alain

Professeur, Université d'Orléans

M MEURENS François

Chargé de recherche/HDR, INRA/Nouzilly

M PFEFFER Sébastien

Chargé de recherche/HDR, CNRS/Strasbourg

M PONCET Didier

Directeur de recherche, INRA/Gif-sur-Yvette

M RASSCHAERT Denis

Professeur, Université François - Rabelais, Tours

REMERCIEMENTS

Je tiens tout d'abord à remercier Sébastien Pfeffer et Alain Legrand pour m'avoir fait l'honneur d'être rapporteurs de ma thèse. Je tiens également à remercier Catherine Dupuy, Didier Poncet et François Meurens pour avoir accepté d'être examinateur de mon travail.

Mon plus grand merci revient à Denis Rasschaert pour m'avoir accepté dans son laboratoire depuis mon master, pour m'avoir appris une quantité phénoménale de points techniques, scientifiques et bien d'autres, tes nombreux conseils et aussi pour m'avoir encadré tout au long de cette thèse même si cela n'a pas dû être facile tous les jours et notamment pour l'écriture de cette thèse. Encore merci.

Je tiens à remercier tout particulièrement Ginette Dambrine pour ses nombreux conseils des plus pertinents, mais aussi pour la relecture de cette thèse. Merci également à Sylvie Laurent pour toutes les fois où je suis venu te demander conseil et où tu m'as aidé et orienté dont notamment pour toute la partie épissage de cette thèse. Merci à Benoit Muylkens, pour mes débuts au laboratoire, pour m'avoir montré les bases pratiques de la biologie moléculaire et pour le travail que l'on a fait ensemble. Merci à vous trois, sans vous je n'aurais pas pu en arriver là.

Je tiens également à remercier tout le reste de l'équipe TLVI ou ex membre de cette fine équipe, Greg, Jennifer, Perrine, Swantje, Elodie, Valérie, Marjolaine, Sylvie, Souheila, Baptiste, Willy et tous les autres et tout particulièrement Thomas et Jérôme pour m'avoir aidé pour l'édition.

Je remercie également Aymeric, Grégory, Nicolas, Audrey et Antoine pour toutes ces années de fac passées ensemble.

Merci à tous mes amis qui m'ont soutenu durant ma thèse dont notamment Julien, Harold, Clément, Alizée, Romuald, Nicolas, mais aussi tous les autres et désolé Clément de t'avoir fait subir ces longues heures d'écriture à mes côtés.

Je remercie également toute ma famille pour leur soutien.

RÉSUMÉ

Le Gallid Herpesvirus 2 (GaHV-2) ou virus de la maladie de Marek (MDV) est un herpesvirus aviaire responsable de l'apparition de lymphomes T chez le poulet. Son génome code pour 25 microARN (miARN) matures regroupés en deux clusters 1 et 2 localisés respectivement au sein des régions répétées inversées TR_L/IR_L et IR_S/TR_S. Les miARN sont de petits ARN d'environ 22 nucléotides, jouant un rôle dans la régulation post-transcriptionnelle par destruction ou séquestration des ARNm et par l'inhibition de la traduction.

Mon sujet de thèse était centré sur l'étude du cluster 1 de miARN de MDV tant au niveau de son implication dans le cycle viral qu'au niveau transcriptionnel. L'étude de la fonctionnalité des deux miARN (miR-M4-5P et miR-M4-3P) issus du mdv1-pré-miR-M4 a débuté classiquement par une recherche de cibles potentielles par bio-informatique. Pour cette étude fonctionnelle nous avons priorisé le mdv1-miR-M4 car les deux brins sont significativement exprimés et le brin 5P est le miARN le plus exprimé durant les différentes phases de l'infection virale (article : Muylkens, Coupeau et coll, Archives of Virology ; 2011). De plus, les sept nucléotides composant la séquence « seed » de mdv1-miR-M4-5P étaient identiques à ceux du miARN cellulaire miR-155, un miARN dont la dérégulation est associée au développement de nombreux lymphomes. Dans un premier temps, j'ai montré en système rapporteur luciférase que les deux miARN orthologues gga-miR-155 et mdv1-miR-M4-5P possédaient des cibles cellulaires communes dont certaines avaient été identifiées pour miR-155 auparavant (CEBPβ, PU.1 et MAP3K7IP2) ou pour la première fois grâce à notre étude (GPM6B, RREB1, c-Myb). Par contre, alors que les deux orthologues ciblaient également potentiellement un gène viral UL28 (DNA packaging terminase subunit 2) nous avons montré par diverses approches complémentaires que seul le mdv1-miR-M4-5P régulait l'expression de ce gène viral spécifique de la phase répllicative du virus. De plus nous avons montré que le miR-M4-3p (différent du miR-155 3P) régulait l'expression d'une autre protéine structurale, UL32 (DNA packaging protein UL32), laissant présager un rôle potentiel dans la balance latence/cycle lytique.

La seconde partie de mon projet de thèse a consisté en l'étude de la régulation transcriptionnelle des miARN du cluster 1 présent dans les régions TR_L et IR_L. D'après la littérature ce cluster semblait scindé en deux parties situées de part et d'autre de l'oncogène *Meq*, avec en amont les miARN : mdv1-miR-M9, M5, M12, M3, M2 et M4 et en aval les miARN : mdv1-miR-M11, M31 et M1. En premier lieu, j'ai réalisé la carte des transcrits de la totalité de la région TR_L/IR_L, par clonage d'amplicons spécifiques et par RACE-PCR. Ainsi, j'ai montré que ces miARN étaient regroupés en réalité en deux clusters. Un premier, regroupant les miARN mdv1-miR-M9 à M4 transcrit soit sous la forme d'un ARNm non épissé recouvrant uniquement ce cluster soit sous une forme intronique au sein de différents grands transcrits épissés. Le second cluster, constitué des miARN mdv1-miR-M11 à M1 pouvait être transcrit soit simultanément au cluster M9-M4 au sein d'un grand intron soit au sein de

deux introns différents suite à divers événements d'épissage de grands transcrits couvrant la totalité de la région TR_L/IR_L. Par ailleurs, les RACE PCR conjuguées à des essais de fonctionnalité suite au clonage de la séquence correspondante en amont du gène rapporteur *luc*, ont permis d'identifier un promoteur juste en amont du cluster M9-M4. En plus d'initier la transcription des clusters de miARN de l'IR_L/TR_L, ce promoteur permettait également la transcription du gène *meq* et de transcrits épissés tels que Meq/v-IL8. Enfin, à l'aide de différentes banques de RACE-PCR correspondants à divers ARNm de la région, j'ai montré que ce promoteur était un promoteur spécifiquement fonctionnel pendant la latence du virus. Le dernier point abordé dans ma thèse concerne la mise en évidence d'édition sur les ARNm issus du brin inverse aux miARN qui était très peu étudié avant mes travaux de thèse. Ainsi, parmi ces transcrits, certains présentaient un phénomène d'hyper-édition recouvrant une région de 2000pb et pouvant toucher plus de la moitié des adénosines présentes. Il faut noter, que l'analyse des clones édités a fait apparaître différentes populations au sein de ces transcrits qui sont en cours d'étude au sein de l'équipe.

Mots clefs : Herpesvirus, GaHV-2, TR_L/IR_L, microARN, épissage, promoteur, édition.

TABLE DES MATIÈRES

REMERCIEMENTS	1
RÉSUMÉ.....	2
TABLE DES MATIÈRES.....	4
LISTE DES TABLEAUX	9
LISTE DES FIGURES	11
ABRÉVIATIONS	14
INTRODUCTION.....	16
1. Régulation transcriptionnelle.....	17
1.1. Synthèse des ARN	17
1.1.1. La transcription.....	17
1.1.2. Transcription par l'ARN Pol II : Les promoteurs et leurs éléments de régulation	17
1.1.2.1. Les facteurs généraux de la transcription	17
1.1.2.2. Transcription basale	21
1.1.2.3. Transcription régulée.....	23
1.1.2.3.1. Les protéines à motifs hélice-tour-hélice (HTH)	25
1.1.2.3.2. Les protéines à motifs en doigts de zinc	25
1.1.2.3.3. Les protéines à domaines basiques	27
1.1.2.3.4. Les protéines à architecture β	27
1.2. Régulations co-transcriptionnelles des ARNm	27
1.2.1. L'adjonction de la coiffe en 5'	27
1.2.2. Addition de la queue polyA en 3'	29
1.2.3. Les mécanismes d'édition.....	31
1.2.3.1. Edition par insertion et délétion de nucléotides	31
1.2.3.2. Edition par conversion de nucléotides.....	32
1.2.3.2.1. Edition par la famille AID/APOBEC.....	34
1.2.3.2.1.1. La protéine APOBEC1	34
1.2.3.2.1.2. La protéine AID	35
1.2.3.2.1.3. Les protéines APOBEC3	35
1.2.3.2.2. Edition par la famille des ADAR.....	35
1.2.3.2.2.1. Les différentes protéines de la famille ADAR.....	35
1.2.3.2.2.2. Mécanisme d'action des protéines ADAR.....	37
1.2.3.2.2.3. Rôles des protéines ADAR.....	39
1.2.4. Le mécanisme d'épissage	42
1.2.4.1. Le spliceosome	44
1.2.4.1.1. Les séquences consensus d'épissage.....	44
1.2.4.1.2. Composition des spliceosomes	46
1.2.4.1.3. Mécanismes d'action des spliceosomes	46

1.2.4.2. L'épissage alternatif	49
1.2.4.2.1. Les séquences régulatrices	49
1.2.4.2.2. L'architecture des exons et des introns	51
1.2.4.2.3. Les structures locales de l'ARNm.....	52
1.2.4.2.4. Régulation de l'épissage par l'ARN Pol II	52
1.2.4.2.5. Les sites d'épissages cryptiques.....	52
1.2.4.3. Devenir des ARNm épissés reconnus par la voie du nonsense mediated decay (NMD)	53
2. Régulation de l'expression génique par les microARN	53
2.1. Généralités	53
2.2. Biogenèse des microARN	55
2.2.1. Transcription des microARN.....	55
2.2.2. Maturation des microARN	55
2.2.2.1. Maturation du pri-microARN.....	57
2.2.2.2. Exportation cytoplasmique du pré-microARN.....	58
2.2.2.3. Maturation cytoplasmique du pré-microARN.....	58
2.2.2.4. Sélection du brin mature du microARN.....	58
2.2.2.5. Formation et régulation du complexe miRISC.....	59
2.3. Interaction microARN/ARNm	60
2.3.1. Reconnaissance de l'ARNm par le microARN	62
2.3.2. Impact de la position du site cible	64
2.3.3. Influence de la structure de l'ARNm.....	64
2.3.4. Influence des protéines de liaison à l'ARN	66
2.4. Mécanismes régulés par les microARN	66
2.4.1. Mécanismes d'inhibition de l'initiation de la traduction	67
2.4.2. Mécanismes d'inhibition post-initiation de la traduction	67
2.4.3. Dégradation de l'ARNm.....	68
2.4.4. Activation de la traduction.....	68
2.5. MicroARN et virus.....	68
2.5.1. Les virus encodant des microARN	68
2.5.2. Régulation du cycle viral par les microARN viraux.....	70
2.5.3. Régulation des réponses cellulaires par des microARN viraux.....	72
2.6. Rôle des ADAR dans l'ARN interférence	73
2.6.1. Régulation de la biogenèse des microARN	73
2.6.2. Régulation de la spécificité des microARN.....	74
3. Le virus de la Maladie de Marek.....	75
3.1. La maladie de Marek.....	75
3.1.1. Historique	75
3.1.2. Transmission.....	75
3.1.3. Les formes cliniques	76

3.1.4. Physiopathologie virale	78
3.2. Le virus de la maladie de Marek	81
3.2.1. Classification	81
3.2.2. Vaccin	83
3.2.3. Structure	85
3.2.4. Organisation génomique de GaHV-2	85
3.2.4.1. Origine de répllication lytique	87
3.2.4.2. Variabilité génique au sein des isolats de GaHV2	89
3.2.4.2.1. Répétitions des 132 pb	89
3.2.4.2.2. Répétitions de 60 pb	90
3.2.5. Transcriptome, protéome et cycle viral de GaHV-2	90
3.2.5.1. Le cycle lytique	91
3.2.5.1.1. Entrée du virus.	93
3.2.5.1.2. Transport de la capsid au noyau et shutt-off cellulaire	93
3.2.5.1.3. Circularisation de l'ADN viral et Transcription des gènes α	95
3.2.5.1.4. Transcription des gènes β	97
3.2.5.1.5. Réplication ADN viral par les gènes β	97
3.2.5.1.6. Expression des gènes γ	97
3.2.5.1.6.1. Les glycoprotéines	98
3.2.5.1.6.2. Les protéines du tégument	98
3.2.5.1.6.3. Les protéines de capsid	99
3.2.5.1.6.4. La lipase virale	99
3.2.5.1.7. Assemblage de la capsid.	99
3.2.5.1.8. Encapsidation.	101
3.2.5.1.9. Sortie du virus.	102
3.2.5.2. Transcriptome des régions TR _L /IR _L et IR _S /TR _S de GaHV-2.	104
3.2.5.2.1. Les microARN de GaHV-2	104
3.2.5.2.2. Transcriptome, protéome de IR _S /TR _S	108
3.2.5.2.3. Transcriptome, protéome de IR _L /TR _L	110
3.2.5.2.3.1. Le gène meq (Marek EcoQ)	110
3.2.5.2.3.2. pp38/pp24	113
3.2.5.2.3.3. La famille des gènes de 1,8 kd	114
3.2.5.2.3.4. Le gène R-LORF5a	115
3.2.5.2.3.5. Le gène R-LORF4	116
3.2.5.2.3.6. Le gène vIL-8	116
3.2.5.2.3.7. vTR	117
3.2.5.2.3.8. Autres ORF présentes dans l'IR _L /TR _L	119
4. Objectifs de travail	122
MATÉRIELS ET MÉTHODES	123

1. Lignées cellulaires et souches virales.....	124
1.1. Lignées cellulaires.....	124
1.2. Souches virales.....	124
1.3. Préparation des CEF et infections virales	124
2. Les vecteurs de clonages et d'expression	125
3. Techniques générales de biologie moléculaire.....	127
3.1. Amplification génique.....	127
3.2. Electrophorèse d'ADN.....	127
3.3. Purification des acides nucléiques	127
3.4. Digestion enzymatique d'ADN.....	128
3.5. Déphosphorylation des vecteurs de clonages.....	128
3.6. Ligature vecteur/insert	128
3.7. Préparation des bactéries électro-compétentes.....	129
3.8. Transformation bactérienne	129
3.9. Criblage des bactéries contenant le plasmide recombinant.....	129
3.10. Midi-préparation d'ADN plasmidique	130
4. Elaboration des différentes constructions.....	130
4.1. Clonage des microARN	130
4.2. Clonage des cibles virales et cellulaires des microARN	130
4.3. Clonage des protéines UL-28 et UL-32 taguées	134
4.4. Clonages des constructions promoteurs du cluster1 de microARN	134
5. Expression des plasmides en système cellulaire eucaryote	136
5.1. Lipofection des cellules adhérentes	136
5.2. Electroporation des cellules en suspension.....	137
5.2.1. Electroporateur EasyjecT Plus.....	137
5.2.2. Electroporateur Amaxa.....	137
5.3. Mesure de l'activité luciférase	138
6. Analyse des ARN totaux	138
6.1. Extraction des ARN totaux	138
6.2. Transcriptase inverse.....	140
6.3. Amplification génique sur ADNc	140
6.4. RACE PCR	140
6.5. Northern -blot.....	144
7. Western-Blot	144
RÉSULTATS	146
1. Etude fonctionnelle de mdv1-miR-M4.....	147
1.1. Comparaison des séquences de mdv1-miR-M4 et gga-miR-155.....	147
1.2. Comparaison du profil d'expression de mdv1-miR-M4 et gga-miR-155.	147
1.3. Recherche de gènes cellulaires ciblés par mdv1-miR-M4-5P et gga-miR-155.	149

1.3.1. Evaluation de la fonctionnalité des miRE lors d'expression ectopique de mdv1-pré-miR-M4 et gga-pré-miR-155.	152
1.3.2. Evaluation de la fonctionnalité des miRE pour mdv1-miR-M4-5P en système endogène.	155
1.4. Analyse de l'effet des deux brins du mdv1-pré-miR-M4 et du gga-pré-miR-155 sur des gènes viraux impliqués dans le clivage et l'encapsidation de l'ADN viral.....	157
1.4.1. Recherche bioinformatique de cibles virales potentielles.....	157
1.4.2. Impact du mdv1-pré-miR-M4 endogène sur les 2 cibles virales	157
1.4.3. Impact du mdv1-pré-miR-M4 et du gga-pré-miR-155 sur les 2 cibles virales lors d'expressions ectopiques.	159
1.4.4. Effet des mdv1-pré-miR-M4 et du gga-pré-miR-155 sur l'expression des protéines UL 28 et UL-32.....	162
2. Etude transcriptionnelle du brin codant les miARN présent sur l'IR _L de GaHV-2.....	162
2.1. Recherche des transcrits codant les miARN présents au sein de l'IR _L de GaHV-2 en latence dans la lignée lymphoïde MSB1.	162
2.1.1. Recherche d'un ARNm spécifique des mdv1-miR-M9 à mdv1-miR-M4.....	166
2.1.2. Recherche de transcrits recouvrant l'ensemble des miARN du cluster 1.	166
2.1.3. Recherche de longs transcrits recouvrant la région IR _L /TR _L de GaHV-2.....	172
2.2. Etude comparative de transcrits encodant les miARN de la région IR _L /TR _L entre les phases de latence et de réplication de GaHV-2.	178
2.3. Etude de la région promotrice des miARN de la région IR _L de GaHV-2.	180
2.3.1. Analyse des éléments <i>cis</i> -régulateurs potentiels de la région promotrice p-miR-M9-M4.....	180
2.3.2. Détermination de la région promotrice minimale de p-miR-M9-M4.	183
2.3.2.1. Construction des formes tronquées du promoteur p-miR-M9-M4.	183
2.3.2.2. Analyse de l'activité luciférase de différentes constructions du promoteur p-miR-M9-M4 en cellules MSB-1.....	185
2.3.2.3. Analyse de l'activité luciférase des différentes constructions du promoteur p-miR-M9-M4 en cellules PA-9.	187
3. Etude transcriptionnelle du brin inverse des miARN de l'IR _L de GaHV-2.....	187
4. Etude de transcrits édités issus des deux brins de l'IR _L de GaHV-2.	190
4.1. Identification de transcrits édités au sein de l'IR _L de GaHV-2.	190
4.2. Etude de la variabilité de l'édition dans la région Edit-29.	192
DISCUSSION	197
RÉFÉRENCES BIBLIOGRAPHIQUES	217
PUBLICATIONS SCIENTIFIQUES	240

LISTE DES TABLEAUX

Tableau 1 : Classification de l'ordre des *Herpesvirales*.

Tableau 2 : Classification du genre *Mardivirus*.

Tableau 3 : Les principales protéines du MDV

Tableau 4 : Expression des 25 microARN matures retrouvés dans les différentes banques issues de la littérature ou des banques obtenues au sein du laboratoire.

Tableau 5 : Oligonucléotides utilisés pour le séquençage des banques et des différentes constructions.

Tableau 6 : Oligonucléotides utilisés pour l'élaboration des vecteurs d'expression des microARN.

Tableau 7 : Oligonucléotides utilisés pour le clonage des cibles sauvages et mutantes des microARN.

Tableau 8 : Elaboration des constructions comportant des mutations internes par mutagenèse dirigée.

Tableau 9 : Oligonucléotides utilisés pour le clonage des protéine taguées UL-28-V5 et UL-32-V5.

Tableau 10 : Oligonucléotides utilisés pour le clonage des construction promotrices de la région p-miR-M9-M4.

Tableau 11 : Oligonucléotides utilisés pour l'amplification du brin inverse aux miARN lors de la reverse transcription.

Tableau 12 : Oligonucléotides utilisés pour l'amplification des banques d'ADNc de la région IRL de GaHV-2.

Tableau 13 : Séquence et position des oligonucléotides utilisés pour lors de la recherche des transcrits édités.

Tableau 14 : Amorces du kit GeneRacer utilisées lors des RACE-PCR.

Tableau 15 : Séquence des oligonucléotides utilisés pour la détection des miARN par Northern blot.

Tableau 16 : Taux d'expression des microARN du cluster 1 de GaHV-2.

Tableau 17: Transcrits identifiés en cellules MSB-1 débutant en amont de mdv1-miR-M9-M4 et permettant d'encoder les miARN du cluster1 de GaHV-2.

Tableau 18: Séquence des bornes donneuses de type U2 présentes sur le brin codant les miARN de l'IR_L de GaHV-2.

Tableau 19: Séquence des bornes accepteurs de type U2 présentes sur le brin codant les miARN de l'IR_L de GaHV-2.

Tableau 20: Séquences nucléotidiques des bornes d'épissage des neuf introns potentiellement de type U12 présent dans l'IR_L de GaHV-2.

Tableau 21 : Localisation des différents sites cis-régulateurs de la région promotrice p-miR-M9-M4.

Tableau 22 : Tableau récapitulatif des différents promoteurs réalisés à partir de la région promotrice p-miR-M9-M4.

Tableau 23 : Séquence des bornes donneuses de type U2 présentes sur le brin inverse des miARN de l'IR_L de GaHV-2.

Tableau 24 : Séquence des bornes accepteuses de type U2 présentes sur le brin inverse des miARN de l'IR_L de GaHV-2.

Tableau 25 : Alignement des miRE ciblés par miR-155 et son orthologue mdv1-miR-M4-5P.

Tableau 26 : Cibles validées pour miR-155 et/ou ses orthologues viraux identifiés chez HHV-8 (KSHV) et GaHV-2.

LISTE DES FIGURES

Figure 1 : Localisation des ER impliqués lors d'une initiation unique de la transcription par l'ARN polymérase II.

Figure 2 : Représentation schématique des deux types d'initiation de la transcription.

Figure 3 : Formation du complexe de pré-initiation de la transcription par l'ARN polymérase II.

Figure 4 : Représentation des 4 grands types de domaines structuraux se liant à l'ADN.

Figure 5 : Changements moléculaires du pré-ARNm liés à l'addition de la coiffe à son extrémité 5'.

Figure 6 : Eléments impliqués dans l'addition de la queue polyA des ARNm.

Figure 7 : Mécanisme général d'édition par insertion et délétion d'uridines.

Figure 8 : Les protéines de la famille AID/APOBEC

Figure 9 : Les protéines de la famille ADAR.

Figure 10 : Altération de l'épissage alternatif d'un transcrit par les protéines ADAR.

Figure 11 : Mécanisme d'édition de l'ARNm du virus de l'hépatite delta.

Figure 12 : Comparaison des séquences régulatrices consensus de type U2 et U12.

Figure 13 : Comparaison des différentes étapes de l'assemblage des spliceosomes de type U2 et U12.

Figure 14 : Représentation schématique des différents types d'épissage alternatif.

Figure 15 : Rôles des protéines SR dans l'épissage.

Figure 16 : Rôles de la protéine hnRNP A1 dans l'épissage.

Figure 17 : Représentation schématique de différentes boucles de régulation entre les microARN et les facteurs de transcription

Figure 18 : Biogenèse des microARN

Figure 19 : Représentation des trois types d'interaction fonctionnelle entre un microARN et sa cible

Figure 20 : Modèle des interactions des différentes régions d'un miARN avec sa cible

Figure 21 : Représentation des différents mécanismes d'inhibition de la traduction par les microARN

Figure 22 : Localisation des microARN au sein des génomes de différents herpesvirus

Figure 23 : Localisation des miARN de HHV-1 au sein du locus LAT

Figure 24 : Chronologie de la pathogénèse d'une souche oncogénique de GaHV-2.

Figure 25 : Micrographie électronique du virion MDV.

Figure 26 : Organisation génomique des herpesvirus des groupes A à F

Figure 27 : Comparaison de la structure génomique des 3 géotypes de Mardivirus, GaHV-2, GaHV-3 et MeHV-1.

Figure 28 : Description et localisation du promoteur bi-directionnel présent dans la région BamHI-H du génome de GaHV-2

Figure 29 : Cascade des gènes de réplication α , β et γ initiée par la protéine VP16 chez HHV-1.

Figure 30 : Comparaison schématique de différentes séquences de VP16.

Figure 31 : Représentation schématique du complexe HCF/Oct1/VP16.

Figure 32 : Assemblage de la capside et encapsidation de l'ADN viral chez HHV-1.

Figure 33 : Localisation et représentation de la structure tige-boucle des 9 pré-microARN issus du cluster 1 de microARN, localisés dans les régions IR_L et TR_L de GaHV-2.

Figure 34 : Localisation et représentation de la structure tige-boucle des 4 pré-microARN issus du cluster 2 de microARN, localisés dans les régions IR_S et TR_S de GaHV-2.

Figure 35 : Alignement des séquences des miR-155 cellulaires humain (hsa) et de *gallus gallus* (gga) avec leurs orthologues viraux HHV-8 (KSHV) et GaHV-2.

Figure 36 : Représentation schématique des trois familles de transcrits alternatifs issus du gène LAT.

Figure 37 : Structure et localisation des différents domaines de la protéine Meq du virus GaHV-2.

Figure 38 : Représentation schématique de l'ensemble des transcrits connus codés par le brin portant l'oncogène *meq* de l'IR_L/TR_L du virus GaHV-2 Md5.

Figure 39 : Structure secondaire de vTR.

Figure 40 : Représentation schématique des deux brins codants de l'IRL du génome de GaHV-2 Md5.

Figure 41 : Carte des vecteurs.

Figure 42 : Elaboration des constructions comportant des mutations internes par mutagenèse dirigée.

Figure 43 : Séquence des oligonucléotides ARN utilisés lors des RACE-PCR.

Figure 44 : Alignement des séquences du gga-pré-miR-155 de *gallus gallus* et de son orthologue viral mdv1-pré-miR-M4 de GaHV-2.

Figure 45 : Détection de mdv1-miR-M4-5P, mdv1-miR-M4-3P, gga-miR-155 et gga-miR-21 par la technique de Northern blot.

Figure 46 : Prédiction des cibles de gga-miR-155 et de son orthologue viral mdv1-miR-M4-5P au sein des 3'UTR de *gallus gallus* par le logiciel TargetScan custom.

Figure 47 : Localisation schématique des miRE au sein de la région 3'UTR de six gènes cellulaires cibles.

Figure 48: Etude comparative de l'effet des microARN mdv1-miR-M4 (miR-4) et gga-miR-155 (miR-155) sur les cibles cellulaires possédant un seul miRE.

Figure 49: Etude comparative de l'effet des microARN mdv1-miR-M4 (miR-4) et gga-miR-155 (miR-155) sur les cibles possédant deux miRE.

Figure 50 : Effet du microARN mdv1-miR-M4-5P endogène sur les 3'UTR des cibles cellulaires présentant un ou deux miRE.

Figure 51: Localisation schématique des miRE de mdv1-miR-M4 présents au sein des régions codantes des gènes viraux UL-28 et UL-32.

Figure 52 : Effet du pré-miARN mdv1-miR-M4 sur les miRE des cibles virales UL-28 et UL-32.

Figure 53 : Evaluation par Western blot de l'effet du mdv1-pré-miR-M4 sur la production des protéines UL-28 et UL-32.

Figure 54 : Représentation schématique du transcrit non épissé encodant les miARN mdv1-miR-M9-M4 du cluster 1 de miARN de GaHV-2 en cellules MSB-1.

Figure 55: Localisation de l'extrémité 3' polyadénylée du transcrit recouvrant les miARN mdv1-miR-M9-M4.

Figure 56: Représentation schématique des transcrits s'initiant en amont de mdv1-miR-M9-M4.

Figure 57: Représentation schématique de l'ensemble des TSS identifiés par 5'RACE PCR dans la région p-miR-M9-M4 en cellules MSB-1.

Figure 58 : Recherche des transcrits spécifiques des gènes des protéines 14kD.

Figure 59: Représentation schématique de l'ensemble des transcrits encodant les miARN du cluster 1 de GaHV-2 en cellule MSB-1.

Figure 60: Comparaison des extrémités 5' des transcrits encodant l'exon 2 de *v-IL8* obtenues par 5' RACE PCR durant la phase de latence ou la réplication de GaHV-2.

Figure 61 : Localisation des différents éléments *cis*-régulateurs présents sur la séquence de la région promotrice p-miR-M9-M4.

Figure 62 : Représentation schématique des différents promoteurs construits à partir de la région promotrice p-miR-M9-M4.

Figure 63 : Etude de l'efficacité transcriptionnelle des promoteurs tronqués de la région promotrice p-miR-M9-M4 en cellules MSB-1.

Figure 64 : Etude de l'efficacité transcriptionnelle des promoteurs tronqués de la région promotrice p-miR-M9-M4 en cellules PA-9.

Figure 65 : Représentation schématique de l'ensemble des transcrits issus du brin inverse au brin codant les miARN du cluster1 de GaHV-2 en cellule MSB-1.

Figure 66 : Représentation schématique des transcrits édités issus de l'IR_L de GaHV-2.

Figure 67 : Alignement de la séquence Edit-185 avec la région correspondante du génome GaHV-2.

Figure 68 : Prédiction des cadres ouverts de lecture de l'exon édité de la séquence Edit-185 en comparaison de la séquence non éditée de GaHV-2 Md5.

Figure 69 : Descriptif des différentes RT-PCR et RT-PCR nichées réalisées dans la région éditée du transcrit Edit-29 en cellules MSB-1.

Figure 70 : Alignement des différentes séquences éditées.

Figure 71 : Alignement des séquences des pré-miR-155 et des pré-miARN de ses orthologues viraux.

Figure 72 : Carte transcriptionnelle du brin codant les miARN de la région IR_L/IR_S de GaHV-2.

Figure 73 : Profil d'épissage dans la région des miARN du virus de Kaposi

Figure 74 : Prédiction de la structure secondaire de l'ARN de la séquence éditée Edit-29 grâce au logiciel mfold.

ABRÉVIATIONS

°C : degré Celsius
μF : micro Farad
ADN : acide désoxyribonucléique
ADNc : acide désoxyribonucléique complémentaire
ARN : acide ribonucléique
ARNm : ARN messenger
miARN : microARN
BET : Bromure d'éthidium
CEF : Chicken Embryo Fibroblast
CMH : Complexe Majeur d'Histocompatibilité
CMV : cytomégalo virus
CaCl₂ : Chlorure de calcium
DNase : DésoxyriboNucléase
dNTP : désoxyribonucléotide tri-phosphate
DTT : dithiothréitol
EBV : Epstein Barr Virus
EDTA : Acide éthylène diamine tétraacétique
g : accélération de l'apesanteur
GaHV : Gallid Herpes Virus
h : heure
HCl : acide chlorhydrique
HHV : Human Herpesvirus
HSV : Hèrpes Simplex Virus
ICP : Infected Cell protein
IE : Immediat Early
IL : Interleukine
IPTG : isopropylthio-β-D-galactoside
IR_L : Internal Repeat Long
IR_S : Internal Repeat Short
KCl : Chlorure de potassium
kb : kilobases
kDa : kilo Daltons
KSHV : Kaposi Sarcoma Herpesvirus
LAT : Latency Associated Transcript
LCMV : Lymphocytic Choriomeningitidis Virus
LMH : Leghorn Male Hépatocyte
LMP1 : Latent Membrane Protein 1
LTR : Long Terminal Repeat
mA : milliampère

mg, µg, ng : milligramme, microgramme, nanogramme
 ml, µl : millilitre, microlitre
 min : minute
 mM, M : millimolaire, molaire
 MDV: Marek's Disease Virus
 MgCl₂ : Chlorure de magnésium
 MgSO₄ : Sulfate de magnésium
 NaCl: Chlorure de sodium
 NaI : Iodure de sodium
 nm : nanomètre
 nt : nucléotide
 ORF : Open Reading Frame
 pb : paire de base
 PBL : Peripheral Blood Lymphocytes
 PBS : Phosphate Buffer Saline
 PCR : Polymerase chain reaction
 pi : post-infection
 PRR : Proline-Rich Repeat
 RISC : RNA Induced Silencing Complex
 RSV : Respiratory Syncytial Virus
 RT-PCR : Reverse Transcriptase PCR
 SDS : Sodium Dodecyl Sulfate
 TAE : Tris/Acétate/EDTA
 TBE : Tris/Borate/EDTA
 TBP : tryptose broth phosphate
 TCR : T Cell Receptor
 TERT : Telomerase Reverse Transcriptase
 TK : Thymidine Kinase
 TNF α : Tumor Necrosis Factor α
 TR : Telomerase RNA
 Tris : tris-(hydroxyméthyl)-aminométhane
 TR_L: Terminal Repeat Long
 TR_S: Terminal Repeat Short
 U_L : Unique Long
 U_S : Unique Short
 v : virulente
 vv : very virulente
 vv+ : very very virulente
 vTR : virus Telomerase RNA
 VZV : Varicella Zoster Virus
 X-Gal : 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside

INTRODUCTION

1. Régulation transcriptionnelle

1.1. Synthèse des ARN

1.1.1. La transcription

Dans les cellules, la transcription correspond à l'étape de synthèse de l'ARN à partir d'une matrice ADN. Lors de l'expression des gènes, la régulation de la transcription est cruciale et déterminante pour leur niveau d'expression, et par conséquent, pour l'ensemble des processus cellulaires. La transcription se déroule en trois étapes, l'initiation, l'élongation et la terminaison.

Chez les eucaryotes, cette réaction est assurée par trois polymérases différentes. L'ARN polymérase I (Pol I), localisée dans le nucléole, est responsable de la synthèse des ARN ribosomaux (ARNr). L'ARN polymérase II (Pol II) synthétise les ARN messagers (ARNm), la majorité des microARN (miARN) ainsi que plusieurs petits ARN nucléaires (snARN : Small Nuclear ARN) dans le nucléoplasme. L'ARN polymérase III (Pol III), fonctionnant en dehors du nucléole, catalyse la synthèse des ARN de transfert (ARNt), de l'ARN ribosomal 5S et de la majeure partie des snARN. Ces ARN polymérases sont assistées par d'autres facteurs en *trans* regroupés sous le terme général de facteur de transcription. Ces derniers, en réponse à un signal donné, sont impliqués dans la reconnaissance des éléments en *cis* contenus dans les séquences promotrices, et ils vont donner le signal pour activer ou réprimer la transcription d'un gène cible, dans un tissu donné.

1.1.2. Transcription par l'ARN Pol II : Les promoteurs et leurs éléments de régulation

L'initiation de la transcription des ARNm requiert une séquence promotrice se situant au niveau de la partie 5' non transcrite des gènes. Cette séquence correspondant à la région où se fixe l'ARN Pol II peut être scindée en deux parties. Une première, appelée promoteur proximal qui constitue le cœur du promoteur et permet l'initiation de la transcription, et une seconde partie qui regroupe les différents éléments *cis*-régulateurs interagissant positivement ou négativement avec le promoteur proximal.

1.1.2.1. Les facteurs généraux de la transcription

Le promoteur proximal recouvre environ les 100 pb situées en amont et les 50 pb en aval du site d'initiation. Il permet le recrutement de la machinerie basale de transcription qui est composée de

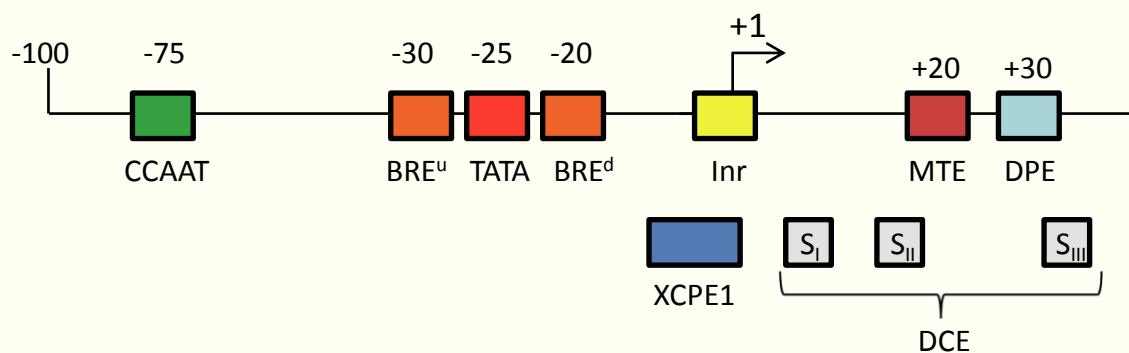


Figure 1 : Localisation des ER impliqués lors d'une initiation unique de la transcription par l'ARN polymérase II. Les différents éléments de réponse sont positionnés par rapport au site d'initiation +1 (d'après Juven-Gershon et Kadonaga., 2010).

l'ARN polymérase II et des facteurs généraux de la transcription (TFIIA, TFIIB, TFIID, TFIIIE, TFIIF, TFIIH) formant le complexe de pré-initiation (PIC : Pre-Initiation Complex). Différents motifs sont généralement retrouvés au sein de cette région (Figure 1) (Juven-Gershon & Kadonaga, 2010).

Le site d'initiation appelé Inr (initiator element) présente une séquence consensus YYANWYY (d'après le code des nucléotides IUPAC), où le A correspond le plus souvent au premier nucléotide de l'ARNm également appelé TSS (Transcription Start site ou +1 de la transcription). La boîte TATA (TATAWAAR) permet le recrutement de la protéine TBP (TATA Binding Protein) et se situe environ 25 à 30 pb en amont de l'Inr. Les séquences BRE (TFIIB-Recognition Element) sont situées de part et d'autre de la boîte TATA, BRE_u (upstream BRE : SSRCGCC) située en amont et BRE_d (downstream BRE : RTDKKKK) en aval de la boîte TATA. De plus, plusieurs sites (DPE, MTE, DCE, S_I à S_{III}) semblent agir en synergie avec l'Inr, dont ils sont séparés par une distance très conservée quelle que soit l'espèce considérée. Le site DPE (Downstream core Promoter Element : RGWYVT) situé environ 30 pb en aval de l'Inr est impliqué dans la reconnaissance de TFIID. Le site MTE (Motif Ten Element : +18CSARCSSAAC+27) agit en synergie entre la boîte TATA et la motif DPE pour le recrutement de TFIID. Les motifs DCE (Downstream Core Element) sont divisés en trois sous éléments : S_I (CTTC) situé en position +6 à +11 du TSS, S_{II} (CTGT) en position +16 à +21 et S_{III} (AGC) en position +30 à +34. Par ailleurs, le motif XCPE1 (X Core Promoter Element 1 : -8DSGYGGRASM+2 chez l'humain), serait présent dans environ 1 % des promoteurs ne présentant généralement pas de boîte TATA.

Bien que ne faisant pas partie du cœur du promoteur, des boîtes CAAT ainsi que des boîtes GC sont retrouvées très fréquemment dans cette région. La boîte CCAAT située entre 60 et 100 pb en amont de l'Inr (Figure 1), possède un fort pouvoir activateur par fixation des facteurs CTF/NF-I (CCAAT-binding Transcription Factor) et de CBF/NF-Y (CCAAT-box Binding Factor). Les boîtes GC (GC box) consistant en des motifs riches en GC permettent la fixation des facteurs ubiquitaires Sp1, Sp3 et de leurs isoformes. Contrairement aux différents *cis*-éléments cités précédemment, ces boîtes sont actives au niveau du promoteur proximal mais également au sein du promoteur distal. Les facteurs Sp1 et Sp3 peuvent recruter et former des complexes avec de nombreuses protéines permettant une activation ou une répression spécifique en fonction du contexte de la région promotrice et du site de liaison. Leur impact est multiple de par leurs interactions avec la machinerie transcriptionnelle de l'ARN Pol II, des protéines de régulation de la transcription ainsi qu'avec la machinerie de remodelage de la chromatine (Davie *et al.*, 2008). Il faut noter que ces facteurs jouent un rôle primordial dans de très nombreux processus cellulaires tels que la régulation de facteurs suppresseurs de tumeurs (p53), de proto-oncogènes (c-myc), de facteurs pro et anti-angiogéniques, pro et anti-apoptotiques, ou encore de facteurs impliqués dans la progression et l'arrêt du cycle cellulaire (Li & Davie, 2010).

Enfin, il n'existe pas de schéma universel pour le contenu des promoteurs proximaux, et bien que la plupart des promoteurs contiennent au moins un de ces éléments en *cis*, aucun ne semble

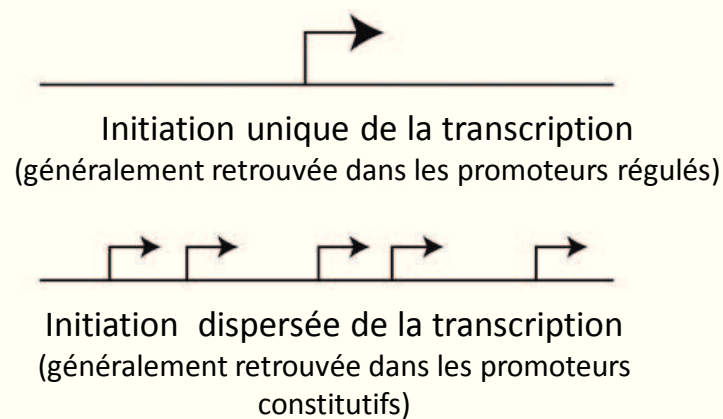


Figure 2 : Représentation schématique des deux types d'initiation de la transcription (d'après Juven-Gershon et Kadonaga., 2010).

Une initiation unique de la transcription est généralement retrouvée au sein des promoteurs régulés qui comportent des *cis*-éléments spécifiques du recrutement des facteurs généraux de la transcription. Pour les initiations de la transcription dispersée, différents sites d'initiation sont retrouvés sur une région allant de 50 à 100 pb. Ce type de promoteur est généralement retrouvé au sein des promoteurs constitutifs ne possédant pas d'éléments généraux *cis*-régulateurs spécifiques du recrutement du complexe d'initiation.

absolument requis. Ainsi, les boîtes TATA ainsi que les séquences BRE, Inr, MTE, DPE et DCE sont essentiellement retrouvées au sein des promoteurs ne présentant qu'un seul site d'initiation de la transcription (FP : Focused Promoter) (Figure 2). Il faut noter que ce type de promoteur à un seul TSS ne représente qu'un tiers des promoteurs retrouvés chez les vertébrés. En effet, la majorité des promoteurs présente une initiation dispersée (DP : Dispersed Promoter) avec potentiellement la présence d'un ou plusieurs sites d'initiation dominants (Figure 2). Par opposition avec les promoteurs FP, ces promoteurs DP sont dépourvus de boîte TATA (TATA less) et de motifs DPE et MTE (Juven-Gershon & Kadonaga, 2010).

1.1.2.2. Transcription basale

Le promoteur minimal est défini *in vitro* comme la plus petite unité nécessaire et suffisante à l'initiation de la transcription et ainsi à l'induction de la transcription basale. *In vivo*, le promoteur minimal pourrait être plus complexe puisque d'autres éléments peuvent être nécessaires.

La formation du complexe de pré-initiation de la transcription se fait en plusieurs étapes et un ordre strict d'assemblage du complexe précède l'initiation de la transcription (Figure 3) (Cooper, 2000). Tout d'abord, le facteur TFIID se lie à l'ADN par l'intermédiaire de la boîte TATA, mais peut également être recruté en absence de cette dernière via l'implication de certains TAF (Transcription Associated Factor) comme TAFIIS qui se lie à l'Inr (Burley & Roeder, 1996). Le recrutement de TFIID crée une courbure de l'ADN qui permet d'une part le rapprochement des séquences situées en amont de la boîte TATA avec celles situées en aval et d'autre part l'ouverture des deux brins d'ADN. Deux voies sont alors possibles : la voie de l'assemblage séquentiel, où les facteurs se fixent dans un ordre établi sur TFIID associé au promoteur, et la voie de l'holoenzyme où tous les facteurs, à l'exception de TFIID, sont pré-assemblés au sein d'un complexe macromoléculaire qui se fixe ensuite sur TFIID associé au promoteur.

Durant l'assemblage séquentiel, TFIIA, qui n'est pas indispensable à la formation du complexe de pré-initiation *in vitro*, permet la stabilisation du complexe TFIID/TATA en interagissant directement avec l'ADN en amont de la boîte TATA. La structure ADN/TFIID/TFIIA est ensuite reconnue par TFIIB (Maldonado *et al.*, 1990). Ce dernier semble impliqué dans la sélection précise du site d'initiation. L'ARN Pol II s'associe alors au complexe conjointement à TFIIF. En effet ces deux complexes sont associés l'un à l'autre pour rejoindre le complexe de pré-initiation en formation (Zawel *et al.*, 1995). Outre son rôle dans l'initiation de la transcription par interaction avec différents partenaires (TFIIB, TFIIE, TBP, ou certaines TAF), TFIIF est impliqué dans la transition initiation-élongation en inhibant un arrêt prématuré de la polymérase (Yan *et al.*, 1999) ainsi que dans l'élongation de la transcription durant laquelle il stabiliserait la polymérase (Lei *et al.*, 1999). L'étape suivante consiste en la liaison de TFIIE avec l'ARN Pol II, TFIIF et TFIID, qui va permettre le

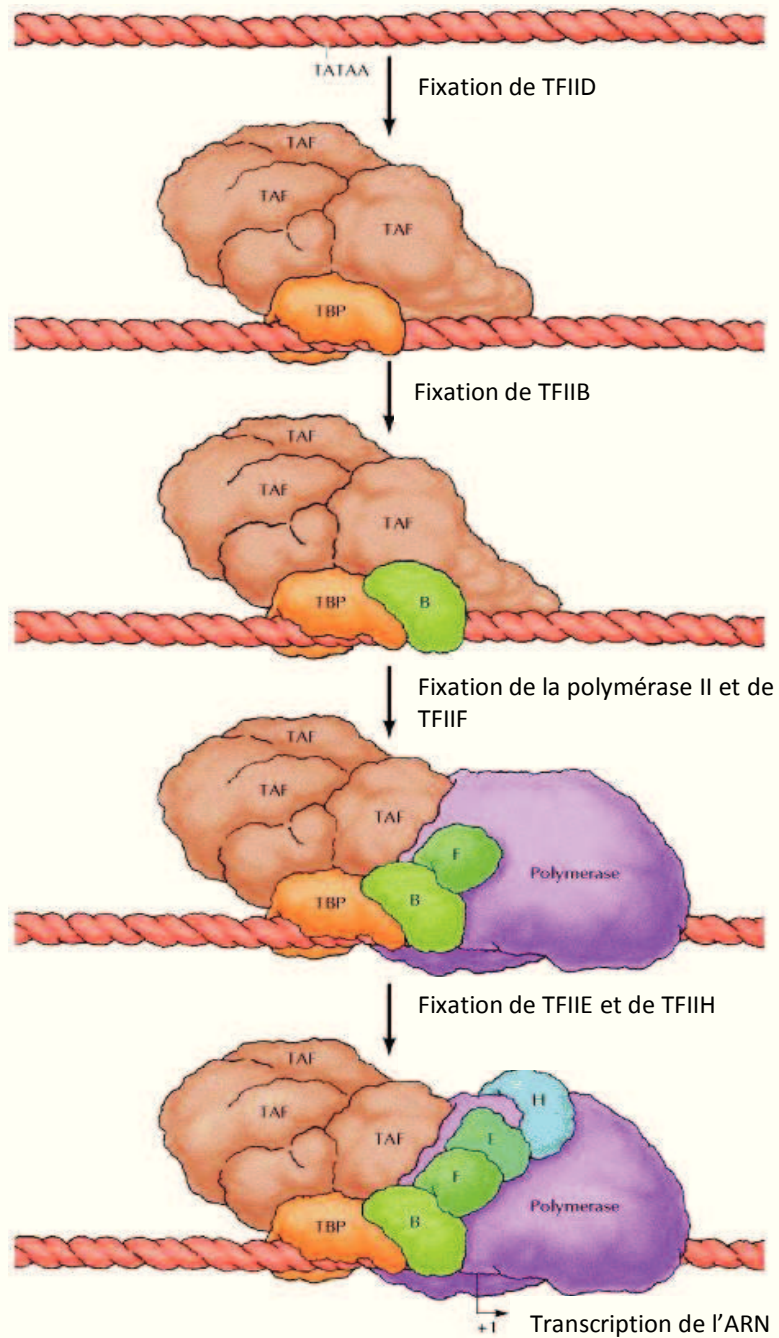


Figure 3 : Formation du complexe de pré-initiation de la transcription par l'ARN polymérase II. Les facteurs essentiels à la formation du complexe *in vitro* se fixent séquentiellement jusqu'à la formation du complexe et l'initiation de la transcription. (d'après Cooper., 2000)

recrutement de TFIIF. Le facteur TFIIF possède deux activités principales : une activité hélicase et une activité kinase. En association avec TFIIE, l'activité hélicase conduit à l'ouverture de la double hélice d'ADN au niveau du promoteur et son activité kinase est responsable de la phosphorylation de la queue C-terminale de l'ARN Pol II (Ohkuma *et al.*, 1995). Cette phosphorylation induisant une modification de la structure tridimensionnelle de la polymérase entraîne la dissociation du complexe d'initiation et le début de la transcription. Enfin, après avoir transcrit une trentaine de bases d'ARN, la Pol II se détache du cœur du promoteur et de la machinerie d'initiation de la transcription et entre en phase d'élongation.

Le modèle de l'holoenzyme permet de concevoir une mobilisation rapide de la machinerie transcriptionnelle. Cependant, dans son acception minimaliste, il ne répond pas aux besoins de flexibilité nécessaires à la spécificité de la réponse aux différents éléments régulateurs. En effet, ce complexe macromoléculaire doit contenir non seulement la machinerie transcriptionnelle (ARN Pol II et facteurs de transcription) mais aussi des protéines régulatrices de la transcription (co-activateurs, complexe médiateur). Aussi au vu de la très grande diversité des promoteurs et des activateurs transcriptionnels, il semble improbable qu'il n'existe qu'un seul et unique type d'holoenzyme. Sachant qu'il a été retrouvé des variations dans la composition de l'holoenzyme selon l'organisme, le tissu étudié ou la méthode de purification, un modèle alternatif a été proposé. Il consiste en l'existence de sous-complexes modulaires et interchangeables, recrutés spécifiquement par les différents régulateurs transcriptionnels lors de l'assemblage de ces divers holoenzymes.

1.1.2.3. Transcription régulée

Le complexe d'initiation composé de l'ARN Pol II et des différents TFII est suffisant pour obtenir une activité transcriptionnelle *in vitro* mais le plus souvent à un faible niveau. L'augmentation de cette activité basale ou sa répression est sous la dépendance de facteurs spécifiques qui vont interagir avec le complexe de pré-initiation. Ces protéines *trans*-activatrices ou *trans*-inhibitrices se lient au promoteur distal, via des séquences *cis*-régulatrices au niveau de l'ADN, appelées « enhancer » lorsqu'ils recrutent des cofacteurs activateurs, ou « silencer » lorsqu'ils recrutent des cofacteurs inhibiteurs. Ces promoteurs distaux peuvent être situés à des milliers de nucléotides du promoteur proximal et agissent sur ce dernier par le jeu de courbures de l'ADN permettant de rapprocher des éléments situés à de grandes distances les uns des autres. De plus, les facteurs de transcription entrent en interaction non seulement avec les séquences spécifiques d'ADN, mais aussi, avec d'autres protéines régulatrices via des domaines d'interaction protéine-protéine ou domaines de dimérisation (homodimère ou hétérodimère).

Quatre grands types de structures protéiques capables de lier l'ADN ont été retrouvés dans les facteurs de transcription (Figure 4), conduisant à la définition de quatre grandes classes de protéines

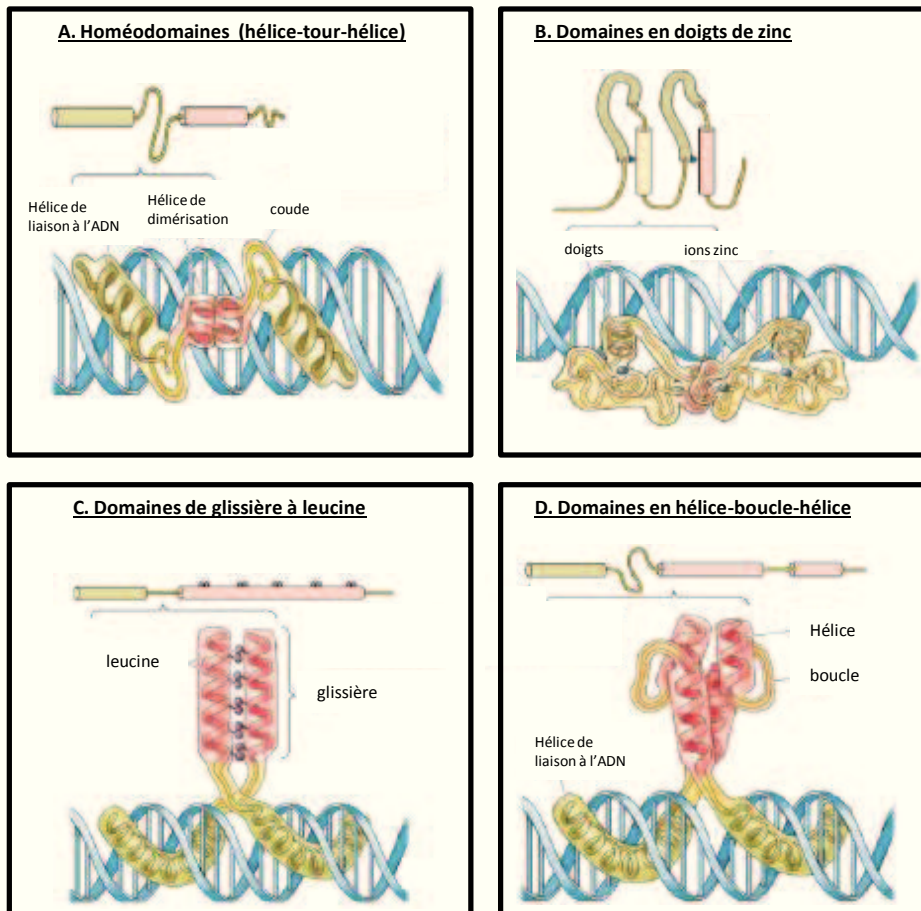


Figure 4 : Représentation des quatre grands types de domaines structuraux se liant à l'ADN. Représentation schématique d'une protéine présentant un homéodomaine (A), des domaines à doigts de zinc (B), un domaine de glissière à leucine (C) et un domaine en hélice-boucle-hélice (D).

(Pabo & Sauer, 1992; Tan & Richmond, 1998). Pour chaque classe de protéine, j'ai restreint les exemples aux protéines rencontrées ultérieurement au cours de ma thèse.

1.1.2.3.1. Les protéines à motifs hélice-tour-hélice (HTH)

Les protéines à motifs hélice-tour-hélice généralement composées de deux hélices α connectées par un tour de quelques résidus (Figure 4A) représentent les protéines se fixant à l'ADN les plus répandues aussi bien chez les procaryotes que les eucaryotes. Ces protéines peuvent être classées en plusieurs groupes structuraux parmi lesquels sont retrouvées les protéines à motifs ailées (E2F-1...), les protéines présentant un homéodomaine (Oct-1, famille des protéines HOX...) ou les protéines ayant un cluster de tryptophane telles que la famille Myb ou PU.1. Ces dernières sont impliquées aussi bien dans la maturation des cellules lymphoïdes que dans l'expression de gènes caractéristiques des lymphocytes T.

Ainsi, les facteurs de transcription de la famille Myb, initialement découverts à partir de la protéine virale v-Myb du virus de myéloblastose aviaire (AMV : Avian Myeloblastosis Virus), sont impliqués dans la transformation de cellules hématopoïétiques. Depuis l'identification de la protéine c-Myb chez le poulet (Gonda & Bishop, 1983), de nombreuses protéines Myb eucaryotes ont été découvertes. Elles présentent toutes un domaine « Myb » très conservé au cours de l'évolution. Ce domaine est constitué de deux ou trois répétitions (R1, R2 et R3) contenant chacune trois résidus tryptophane légèrement espacés. Ces protéines sont essentielles à la survie de la cellule, elles sont impliquées dans la régulation du cycle cellulaire et de la différenciation. Le proto-oncogène c-Myb est impliqué, par exemple, dans le contrôle de la prolifération lymphocytaire, notamment lors de l'expression de gènes de différenciation des cellules T tels que c-Kit, CD4, TCR γ , TCR δ (Allen *et al.*, 2001; Bender *et al.*, 2004; Hernandez-Munain & Krangel, 2002; Hsiang *et al.*, 1995).

Le facteur de transcription PU.1, appartenant à la famille des facteurs Ets, est un régulateur clef du développement précoce de toutes les cellules hématopoïétiques (Kastner & Chan, 2008) ; notamment il permet, sous le contrôle du miARN cellulaire miR-155, de moduler la réponse immunitaire des lymphocytes B (Vigorito *et al.*, 2007). De plus, il semble que sa dérégulation participe à la transformation des lignées myéloïdes (Kastner & Chan, 2008).

1.1.2.3.2. Les protéines à motifs en doigts de zinc

Les motifs des protéines à motifs en doigts de zinc forment de petits domaines, repliés indépendamment du reste de la structure protéique autour d'un ion zinc qui stabilise la formation d'un feuillet β contre une hélice α (Tan & Richmond, 1998). Parmi ces protéines, on retrouve de nombreux facteurs indispensables à la cellule tels que Sp1, TFIIA ou le facteur GATA 3 qui est un régulateur important du développement des lymphocytes T (Chou *et al.*, 2010).

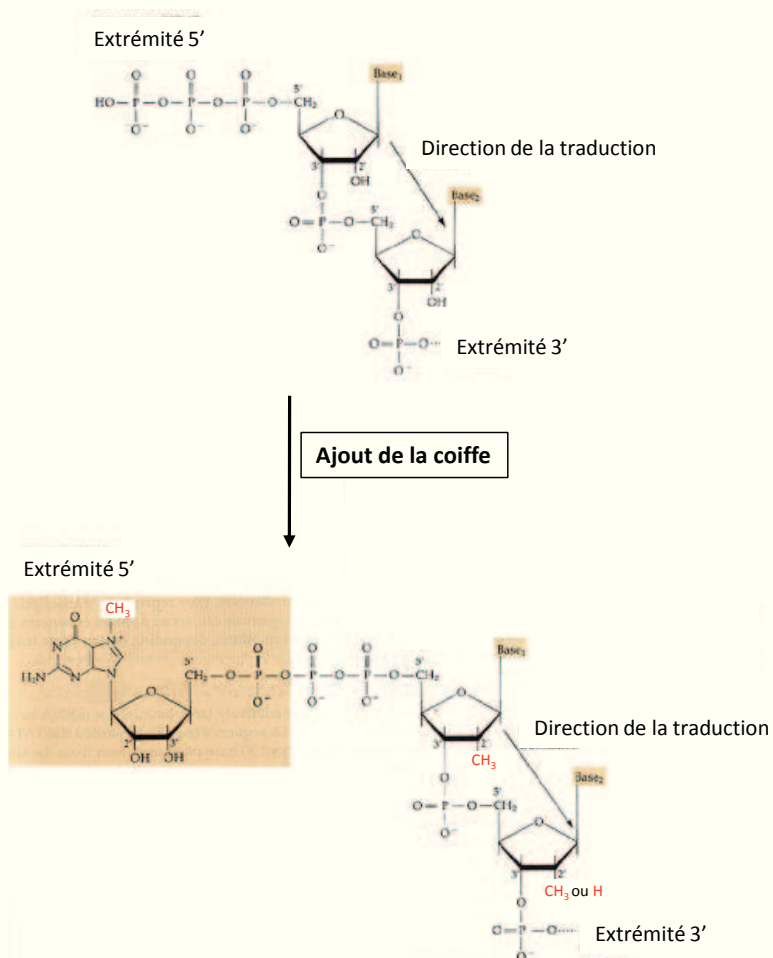


Figure 5 : Changements moléculaires du pré-ARNm liés à l'addition de la coiffe à son extrémité 5'. Les résidus méthylés après l'adjonction du nucléotide à guanine sont représentés en rouge.

1.1.2.3.3. Les protéines à domaines basiques

Elles regroupent les protéines à domaines en hélice-boucle-hélice (bHLH) caractérisés par une hélice α amphipathique, une boucle puis une autre hélice α amphipathique (Figure 4D) (Murre *et al.*, 1989) et les protéines à domaines de glissière à leucine (leucine zipper ou bZIP). Ces dernières sont caractérisées par une répétition de plusieurs leucines (Figure 4C) (Landschulz *et al.*, 1988), permettant la formation d'homodimères ou d'hétérodimères avec d'autres protéines possédant également un domaine bZIP. On retrouve parmi cette famille de protéines les composants AP1-like (Fos, Jun, Maf...) ou les facteurs C/EBP-like (C/EBP α , C/EBP β ...). De plus, certains facteurs de transcription peuvent combiner ces deux types de domaine comme par exemple Myc ou Mad/Max qui possèdent un motif bZIP ainsi qu'un motif bHLH.

1.1.2.3.4. Les protéines à architecture β .

Il est difficile de trouver une caractéristique commune à tous les facteurs composant cette superclasse tant leur architecture β n'est pas partagée par tous. Ainsi, parmi ces différentes classes, on retrouve la protéine p53 qui est un facteur clef du contrôle du cycle cellulaire et de l'apoptose, ainsi que des facteurs de transcription activés en réponse à la stimulation par les cytokines tels que STAT5, NF- κ B, NF-AT. Ces derniers sont particulièrement impliqués dans la survie et la prolifération des lymphocytes T.

1.2. Régulations co-transcriptionnelles des ARNm

Au sein des cellules eucaryotes, la grande majorité des ARN pré-messagers (pré-ARNm) transcrits par l'ARN polymérase II subissent une série de modifications co-transcriptionnelles permettant leur maturation : ajout de la coiffe en 5', épissage des introns, édition et poly-adénylation en 3'.

1.2.1. L'adjonction de la coiffe en 5'

L'adjonction de la coiffe (ou cap) a lieu dès le début de la transcription avant que la chaîne ne compte plus de 30 nucléotides. Elle consiste en l'adjonction d'un nucléotide à guanine sur l'extrémité 5' non traduite de l'ARNm suivie de sa méthylation sur l'azote 7 de la base, ainsi que parfois la méthylation en 2' du ribose du premier et second nucléotide du transcrit primaire (Figure 5). Le phosphate γ initialement présent à l'extrémité 5' du pré-ARNm est hydrolysé par l'ARN

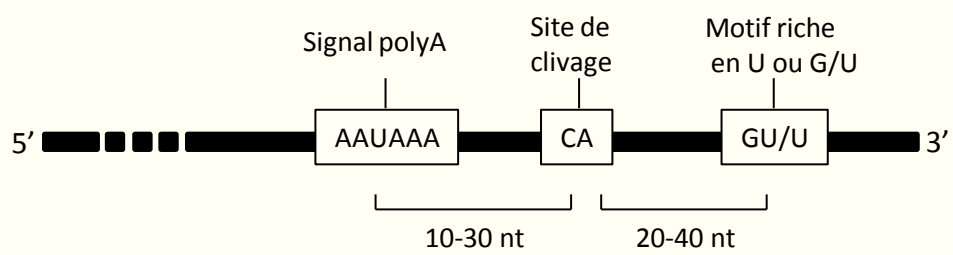


Figure 6 : Eléments impliqués dans l'addition de la queue polyA des ARNm.

triphosphatase (RTase). Ensuite, l'ARN guanyltransférase (GTase) transforme une molécule de GTP en GMP pour la transférer à l'extrémité diphosphorylée du pré-ARNm. Il y a alors formation d'un pont 5'-5' triphosphate, suivie de la méthylation par l'ARN guanine-7-méthyltransférase (MTase) ainsi que par la 2'-O-méthyltransférase. Cette étape est un processus directement couplé à la transcription puisque la GTase et la MTase sont toutes deux directement recrutées par l'extrémité carboxyterminale de l'ARN Pol II lorsque cette dernière est phosphorylée.

Il en résulte que l'extrémité 5' de l'ARNm n'est pas porteuse des trois acides phosphoriques libres habituels, ce qui permet la protection de l'extrémité 5' de l'ARNm en limitant sa reconnaissance par les 5' → 3' exonucléases. La coiffe est également nécessaire à l'exportation de l'ARNm vers le cytoplasme et à la liaison de ce dernier avec la petite sous-unité du ribosome lors de l'étape d'initiation de la traduction. De plus, elle est indispensable à la poursuite des étapes ultérieures de maturation des pré-ARNm par l'intermédiaire du complexe CBP20/CBP80 (Cap Binding Protein 20 et 80). En effet, le recrutement de ce complexe va participer à la mise en place de la machinerie d'épissage du premier intron et va influencer l'efficacité de la maturation en 3' (Proudfoot *et al.*, 2002).

1.2.2. Addition de la queue polyA en 3'

A l'exception des transcrits codants pour les histones, tous les messagers eucaryotes comportent une queue polyadénosine (polyA) à leur extrémité 3'. Le site de polyadénylation est codé au niveau du gène. Il correspond à une séquence ubiquitaire AAUAAA située 20 à 30 nucléotides en amont du site de clivage CA et par une séquence riche en GU ou U située une trentaine de nucléotides en aval de ce site de clivage (Figure 6).

Il a été identifié une dizaine de variants du site de polyadénylation chez l'humain parmi lesquels, les deux séquences AAUAAA et AUUAAA sont majoritairement retrouvées (Beaudoing *et al.*, 2000). Ces séquences sont reconnues spécifiquement par un complexe protéique appelé CPSF (Cleavage and Polyadenylation Specific Factor), tandis que la séquence riche en GU est reconnue par un complexe CstF (Cleavage Stimulation Factor). L'association de ces deux complexes avec d'autres composants, comme l'ARN polymérase II et la poly(A) polymérase (PAP), conduit à la coupure de la molécule de pré-ARNm au niveau du site CA. La PAP qui est une terminal transférase n'utilisant pas de matrice ADN, ajoute ensuite environ 200 adénosines, sous le contrôle d'une poly(A) binding protein II (PABP2). En effet, cette dernière active la PAP et par la suite contrôle la longueur de la queue poly(A). Il faut noter qu'en plus d'être nécessaire à l'initiation de la traduction, cette queue poly(A) confère de la stabilité à l'ARNm et sa longueur diminue progressivement au fur et à mesure des cycles de traduction du message.

La maturation alternative en 3' consiste au choix entre deux ou plusieurs sites de clivage et de polyadénylation au sein du transcrit primaire et il en résulte une diversité de la région 3' non traduite

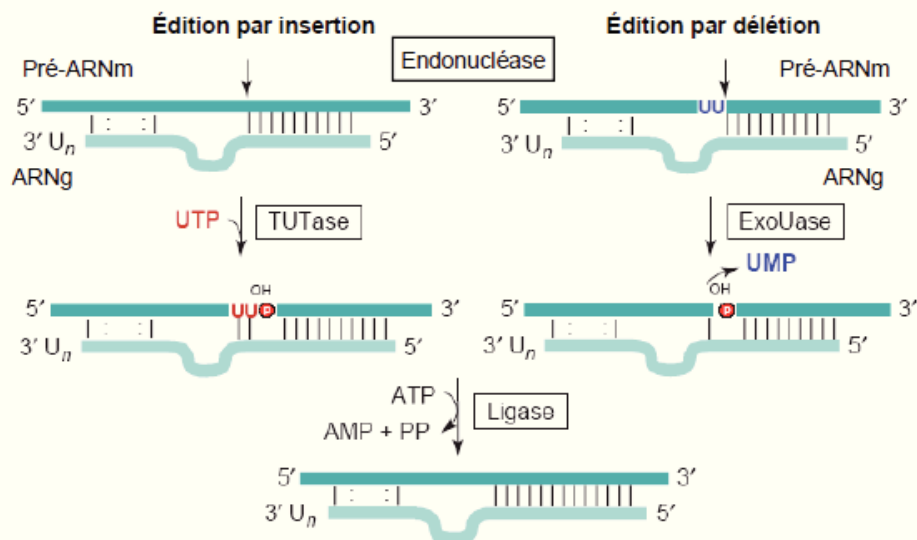


Figure 7 : Mécanisme général d'édition par insertion et délétion d'uridines. Ce mécanisme d'édition a été retrouvé dans les ARNm mitochondriaux des trypanosomes (d'après Stuart et al., 2005).

(3'UTR) ou de la phase codante d'un même pré-ARNm (Edwards-Gilbert *et al.*, 1997). Cette variabilité de longueur de la région 3' UTR d'un messager a une incidence sur la stabilité des messagers et sur la régulation post-transcriptionnelle par les miARN, les 3' UTR les plus courtes pouvant perdre des cibles de régulation par certains miARN. Un modèle suggère également que les composants de la machinerie d'épissage interagissent avec ceux impliqués dans la maturation en 3' définissant ainsi très tôt lors de l'épissage, l'exon 3' terminal du pré-ARNm (Berget, 1995).

1.2.3. Les mécanismes d'édition

Le terme d'édition décrit l'ensemble des processus induisant des modifications ponctuelles de nucléotides au sein de la séquence ADN ou ARN conduisant à un transcrit mature dont la séquence diffère de la séquence génomique ARN ou ADN.

Alors que l'édition de l'ADN via des protéines de la famille APOBEC a été découverte très récemment (Harris *et al.*, 2002), l'édition de l'ARN a été décrite pour la première fois en 1986, dans les mitochondries des trypanosomes (Benne *et al.*, 1986). Par la suite ce mécanisme a été retrouvé dans de nombreuses classes d'organismes (mammifères, plantes, champignons, bactéries ainsi que certains virus). L'édition des ARNm semble restreinte aux eucaryotes, tandis que l'édition des ARN non codants (ARNnc), en particulier des ARNt, semble commune aux procaryotes et aux eucaryotes. Par ailleurs, le terme d'édition recouvre 2 grands mécanismes : l'insertion et la délétion de nucléotides ou la conversion de nucléotides.

1.2.3.1. Edition par insertion et délétion de nucléotides

L'édition par insertion ou délétion de nucléotides est un phénomène plus rare que l'édition par conversion de nucléotides. Ce phénomène a été principalement caractérisé chez les trypanosomes mais il a également été décrit chez d'autres organismes tels que les myxomycètes, ou certains virus.

L'édition retrouvée au sein de très nombreux ARNm mitochondriaux des trypanosomes consiste majoritairement en l'insertion d'uridines et en plus faible proportion en leur délétion. Ce phénomène d'édition peut avoir des conséquences importantes comme par exemple pour le gène *cox3* (subunit 3 of the cytochrome *c* oxydase) où environ 90% des codons sont générés par édition (Horton & Landweber, 2002). En effet, l'ARNm de *cox3* long de 463 nucléotides après transcription subit une insertion de 547 uridines et une délétion de 41 uridines pour donner l'ARNm mature d'une longueur de 969 nucléotides qui permet la traduction du polypeptide conservé COX3 (Brennicke *et al.*, 1999). L'insertion et la délétion d'uridines se déroule en plusieurs étapes (figure 7) (Stuart *et al.*, 2005). Tout d'abord, un ARN guide, qui forme une classe unique d'ARNnc, s'hybride par sa région

complémentaire de la région en aval du site d'édition au pré-ARNm. Après coupure par une endonucléase, les uridines sont ajoutées par une TUTase (Terminal Uridyl Transferase) dans le sens 3'→5' en utilisant l'ARN guide comme matrice, ou délétées dans le sens 3'→5' par une ExoUase (U-specific exonucléase). Enfin, lors de la dernière étape, l'ARNm édité est religaturé (Figure 7). Il faut noter que le profil d'édition varie au cours du cycle du parasite, laissant supposer un rôle de l'édition dans la régulation de la production d'énergie par la chaîne respiratoire. De plus, l'insertion ou la délétion d'uridines n'est pas le seul phénomène d'édition chez les trypanosomes et des conversions C→U ont également été retrouvées sur un ARNt (Alfonzo *et al.*, 1999) et sur l'ARN 7SL (Ben-Shlomo *et al.*, 1999), permettant à ces ARNnc transcrits dans le noyau d'être fonctionnels.

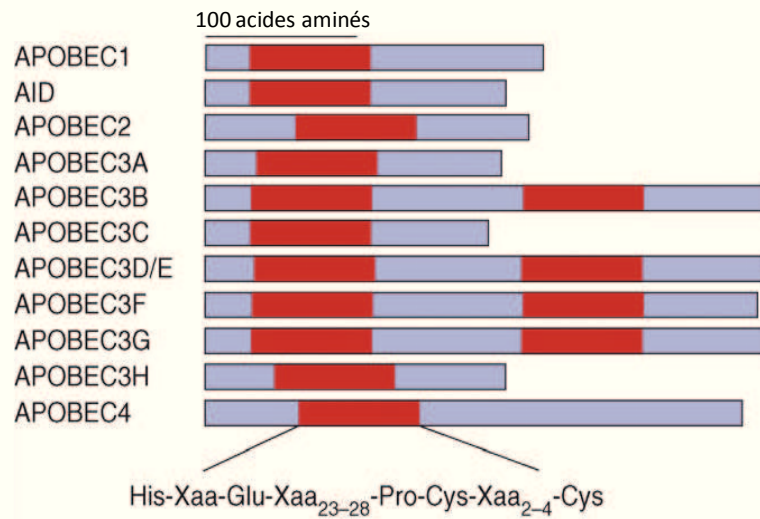
L'édition chez certains myxomycètes consiste en un mélange de plusieurs processus complexes et très fréquemment retrouvés au sein des ARNm codés par le génome mitochondrial de ces organismes. Bien que peu caractérisés, de nombreux types d'édérations comme l'insertion d'un des quatre nucléotides A, G, C ou U sous forme de monomères ou de dimères, des phénomènes de délétions et des conversions de nucléotides ont été décrits. Enfin, ce qui semble unique dans le monde des eucaryotes, certains transcrits sont régulés par quatre phénomènes différents d'édition (Horton & Landweber, 2000), comme par exemple le gène *cox1* (cytochrome c oxydase subunit 1), dont le pré-ARNm subit une insertion de 59 cytosines, 1 uridine et de 3 dinucléotides différents, ainsi que quatre conversions C→U afin de pouvoir coder la protéine fonctionnelle (Gott *et al.*, 1993).

Pour finir, un phénomène d'édition co-transcriptionnel des ARN lié au « bégaiement » de l'ARN polymérase a également été mis en évidence chez certains virus à ARN simple brin négatif. Ce mécanisme correspond à l'arrêt et au glissement du complexe de transcription sur la matrice. Ainsi, le gène de la phosphoprotéine P chez les paramyxovirus permet d'encoder quatre protéines différentes (P, V, W et C) grâce à l'adjonction par la polymérase d'une séquence de 1 à 11 G au niveau d'un site spécifique de l'ARNm, permettant la modification du cadre de lecture et la traduction des protéines V et W (Kulkarni *et al.*, 2009). De plus, un phénomène semblable mais avec l'addition d'une adénine a également été retrouvé au niveau d'un site précis des transcrits de la glycoprotéine du virus Ebola, aboutissant à deux glycoprotéines fonctionnelles différentes (Sanchez *et al.*, 1996; Volchkov *et al.*, 1995).

1.2.3.2. Edition par conversion de nucléotides

La conversion de nucléotides est retrouvée dans l'ensemble des organismes. D'une manière générale, les enzymes capables d'édition par conversion de nucléotides jouent un rôle très important dans de nombreuses voies métaboliques, de la bactérie jusqu'à l'homme. Chez les animaux, un nombre restreint de sites d'édition ont été identifiés dans les régions codantes des gènes, cependant l'édition semble être plus massive dans les régions non codantes des ARNm (5' et 3' UTR) où elle

A.



B.

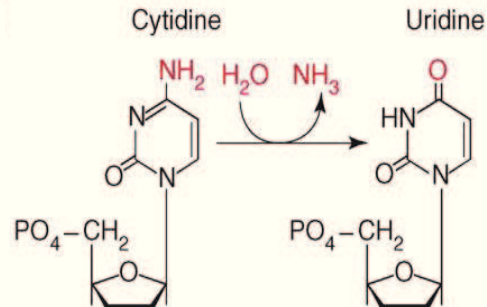


Figure 8 : Les protéines de la famille AID/APOBEC (d'après Holmes et al., 2007).

(A) Organisation des différents domaines des protéines APOBEC humaines. Les motifs « cytidine désaminase » correspondant à la séquence consensus indiquée en partie inférieure du schéma sont figurés par un rectangle rouge au sein des différentes protéines (en bleu). (B) Désamination d'une cytosine en uridine catalysée par les enzymes AID/APOBEC.

joue un rôle fondamental dans des processus de régulation post-transcriptionnelle, de maintien de l'intégrité génomique, mais aussi de défense antivirale. Deux types majeurs de conversions issus de deux familles de protéines ont été identifiés : la désamination C en U d'ARN ou d'ADN simple brin par la famille de protéine AID/APOBEC (Activation-Induced cytidine Deaminase/Apolipoprotein B mRNA Editing enzyme, Catalytic polypeptide-like) et la désamination A en I (Inosine) d'ARN double brin par la famille de protéines ADAR (Adenosine Deaminases Acting on RNA).

1.2.3.2.1. Edition par la famille AID/APOBEC

Une dizaine de protéines ont été identifiées au sein de cette famille de cytidine désaminases (Figure 8A). Elles permettent la conversion d'une cytidine en uridine sur des matrices ADN et ARN (Figure 8B). Bien que la capacité de cette famille de protéines à modifier l'ADN constitue un danger potentiel pour la cellule du fait de son action mutagène, elles jouent un rôle important dans les systèmes immunitaires inné et adaptatif (Prochnow *et al.*, 2009), et de par l'action d'APOBEC1, un rôle clef dans le métabolisme lipidique. Parmi les protéines peu caractérisées de cette famille, la protéine APOBEC2 semble spécifiquement exprimée dans les muscles squelettiques et cardiaques (Liao *et al.*, 1999), cependant son rôle reste inconnu car elle ne semble pas présenter d'activité cytosine désaminase et aucun effet phénotypique n'est détecté chez des souris invalidées pour *apobec2* (Mikl *et al.*, 2005). La protéine APOBEC4, qui a été prédite *in silico*, semble plutôt exprimée dans les testicules des mammifères, cependant aucun rôle ni même sa présence *in vivo* n'ont été démontrés (Rogozin *et al.*, 2005).

1.2.3.2.1.1. La protéine APOBEC1

L'exemple probablement le plus connu d'édition d'ARNm chez les mammifères concerne l'apolipoprotéine B (ApoB). Ce mécanisme d'édition spécifique de tissu, découvert en 1987, permet la synthèse de deux protéines ApoB100 (512 kDa) et ApoB48 (241 kDa) à partir d'un même gène (Chen *et al.*, 1987; Powell *et al.*, 1987). Dans le foie, le transcrit dirige la synthèse d'ApoB100, tandis que celui-ci va être spécifiquement édité dans l'intestin par la protéine APOBEC1. Ainsi, l'édition de la cytosine présente en position 6666 va induire la modification d'un codon glutamine (CAA) en codon stop (UAA) et permettre la traduction de la protéine ApoB48.

Bien qu'indispensable, APOBEC1 n'est pas suffisante pour induire seule l'édition, et nécessite *in vitro* au moins une autre protéine connue sous le nom d'ACF (APOBEC1 Complementary Factor) (Mehta *et al.*, 2000). Cette protéine se lie à APOBEC1 et à la région qui encadre le site d'édition, permettant de modifier la structure de l'ARNm et de diriger la sous unité catalytique APOBEC1 sur le site à éditer (Chester *et al.*, 2004; Maris *et al.*, 2005). Cependant, en dehors de la caractérisation de ce complexe minimal d'édition, l'identité et le rôle des autres facteurs formant le complexe d'éditosome reste mal connu. Il faut noter que la surexpression de la protéine APOBEC1 peut induire un

phénomène d'hyper-édition sur certains transcrits cellulaires dont l'apolipoprotéine B (Petit *et al.*, 2009; Yamanaka *et al.*, 1997). Par ailleurs, il a été montré que cette protéine est aussi capable d'induire ce phénomène d'hyper-édition *in vivo* sur de l'ADN simple brin et sur l'ARN génomique du gammarétrovirus murin Friend-murine leukemia virus (Petit *et al.*, 2009).

1.2.3.2.1.2. La protéine AID

La protéine AID (Activation-Induced cytidine Deaminase) de 24 kDa, qui est exprimée spécifiquement dans les lymphocytes B des vertébrés (Muramatsu *et al.*, 1999), est nécessaire à la production et la maturation des anticorps. Cette protéine ciblant l'ADN simple brin joue un rôle central dans les deux mécanismes nécessaires à la production d'anticorps ayant une haute affinité, qui sont l'hyper-mutation somatique et la commutation isotypique de classe des immunoglobulines (Muramatsu *et al.*, 2000).

1.2.3.2.1.3. Les protéines APOBEC3

Cette sous-famille spécifique des mammifères est constituée de 7 protéines (APOBEC3 A à H), dont la plupart permettent une inhibition de la réplication de divers rétrovirus ainsi que d'autres rétroéléments endogènes. Quatre de ces membres (APOBEC3B, D/E, F et G) sont potentiellement capables d'inhiber le virus de l'immunodéficience humaine (Prochnow *et al.*, 2009). Cependant, la protéine Vif (Viral Infectivity Factor) du VIH-1 permet de contourner cette défense antivirale en séquestrant APOBEC3G et en l'entraînant vers le protéasome par ubiquitinylation (Conticello, 2008). Dans le cas d'une infection avec un virus délété du gène *vif*, la protéine APOBEC3G est encapsidée dans les particules virales. L'édition s'opère sur l'ADN négatif suite à la transcription inverse. L'enzyme APOBEC3G induisant de nombreuses désaminations de cytidine en uridine aurait ainsi une action antivirale en bloquant la réplication du virus ou en générant des génomes édités moins infectieux. Le virus de l'hépatite B (VHB), qui utilise une transcriptase inverse durant son cycle, est également la cible de certaines protéines APOBEC3. En effet, ces dernières semblent inhiber l'encapsidation de l'ARN pré-génomique et ainsi le pouvoir infectieux et la réplication du virus (Holmes *et al.*, 2007). En plus de leur pouvoir hypermutagène, les protéines APOBEC3 auraient la capacité de bloquer la mobilité des rétrotransposons à l'aide d'un mécanisme différent et encore inconnu, ne faisant pas intervenir l'activité d'édition (Holmes *et al.*, 2007).

1.2.3.2.2. Edition par la famille des ADAR

1.2.3.2.2.1. Les différentes protéines de la famille ADAR

Le phénomène d'édition par désamination d'une adénine en inosine (A→I) a été décrit pour la première fois chez *Xenopus laevis* (Bass & Weintraub, 1987; Rebagliati & Melton, 1987). Elle est

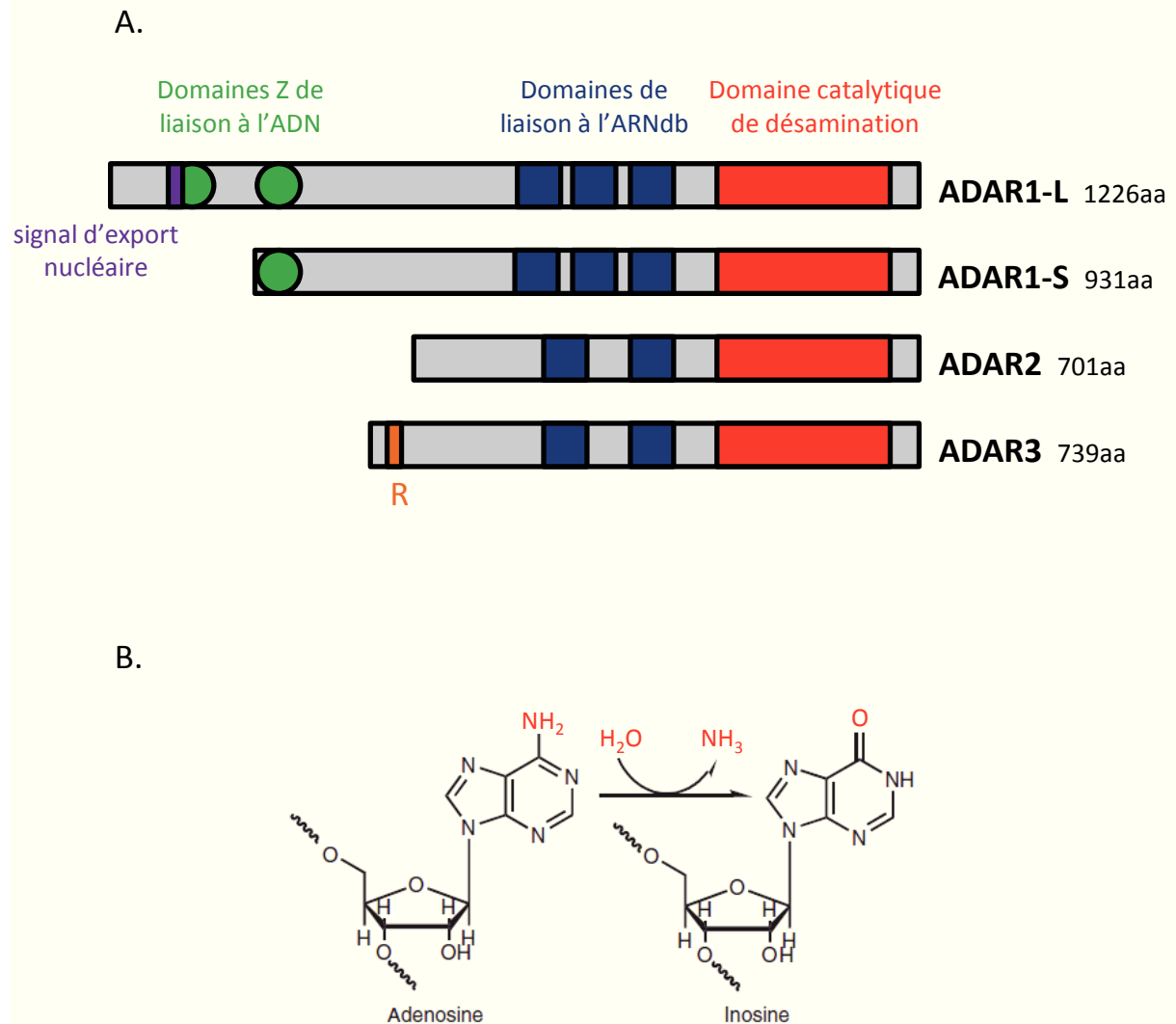


Figure 9 : Les protéines de la famille ADAR.

(A) Organisation des différents domaines des protéines ADAR humaines. Les domaines catalytiques à activité désaminase sont indiqués en rouge, les motifs de liaison à l'ARN double brin dsRBD (dsRNA-Binding Domain) sont indiqués en bleu, les domaines Z de liaison à l'ADN (Z DNA Domain) sont représentés par un disque vert, le signal d'export nucléaire présent chez ADAR1-L est indiqué en violet et le site R d'ADAR3 (riche en Arginine) correspondant à un domaine de liaison à l'ARN simple brin est représenté en orange. (B) Désamination d'une adénosine en inosine catalysée par les enzymes ADAR (d'après Zinshteyn et al., 2010).

catalysée par la famille des protéines ADAR (Figure 9), qui comprend les quatre protéines ADAR1-L, ADAR1-S, ADAR2 et ADAR3. Ces protéines sont conservées chez tous les vertébrés (Figure 9), et des membres de cette famille ont également été retrouvés chez l'ensemble des métazoaires.

ADAR1-S, ADAR1-L et ADAR2 sont exprimées dans tous les tissus tandis qu'ADAR3 est exprimée spécifiquement dans le cerveau. Les deux formes de la protéine ADAR1 spécifiques des vertébrés (ADAR1-S, ADAR1-L) sont traduites à partir de deux messagers emboîtés, issus de l'utilisation de deux promoteurs différents. Il faut noter que le messenger d'ADAR1-L qui est transcrit sous la dépendance d'un promoteur régulé par l'interféron est produit lors d'infections virales. De plus, contrairement à toutes les autres protéines ADAR, ADAR1-L possède dans sa partie NH2 terminale non présente dans ADAR1-S un signal d'export nucléaire qui lui permet d'agir dans le cytoplasme, siége de la réplication de beaucoup de virus à ARN (Figure 9) (Poulsen *et al.*, 2001).

Tous les membres de cette famille contiennent un domaine catalytique conservé dans leur partie COOH terminale (Figure 9) ainsi que deux ou trois domaines de liaison à l'ARN double brin dsRBD (Double Strand RNA Binding Domain) leur permettant de cibler avec une forte affinité les ARN double brin (ARNdb) ainsi que les structures secondaires double brin des ARN simple brin (Nishikura, 2010).

Par ailleurs, il existe une autre famille de protéines responsable de la désamination A→I, la famille des ADAT (Adenosine Deaminase acting on tRNA), qui est retrouvée conservée dans l'ensemble des eucaryotes et dont un orthologue a également été identifié chez différentes bactéries comme *Escherichia coli* (Wolf *et al.*, 2002). Cette famille de protéines est responsable de l'édition des ARNt sur ou à proximité de la séquence de l'anticodon (Nishikura, 2010).

1.2.3.2.2. Mécanisme d'action des protéines ADAR

De par leur activité de désamination A→I, les protéines ADAR modifient l'information codée par le génome ou la structure de l'ARN substrat. En effet, lorsqu'un résidu inosine est incorporé dans une séquence codante d'un ARNm, il va être interprété comme une guanosine par la machinerie de traduction, ce qui peut altérer la nature du codon spécifié et par là-même la séquence protéique. De plus, l'édition peut avoir un impact direct sur l'ARNm en modifiant son épissage via différents mécanismes (Figure 10) ou en modifiant les structures secondaires de cet ARN. Ainsi, elle peut déstabiliser la structure double brin de l'ARN en remplaçant un appariement A:U par un mésappariement IxU ou, au contraire, augmenter la stabilité d'une tige par remplacement d'un mésappariement AxC par une paire I:C (Levanon *et al.*, 2004).

Historiquement, les premières cibles identifiées pour les protéines ADAR ont été retrouvées dans les régions codantes, cependant, notamment dans le cerveau, il semblerait que la majorité des sites d'édition se situent dans les régions non codantes des ARNm, telles que les introns et les régions non traduites 3' et 5' (Hundley & Bass, 2010). De plus, les protéines ADAR ciblent les ARNnc tels

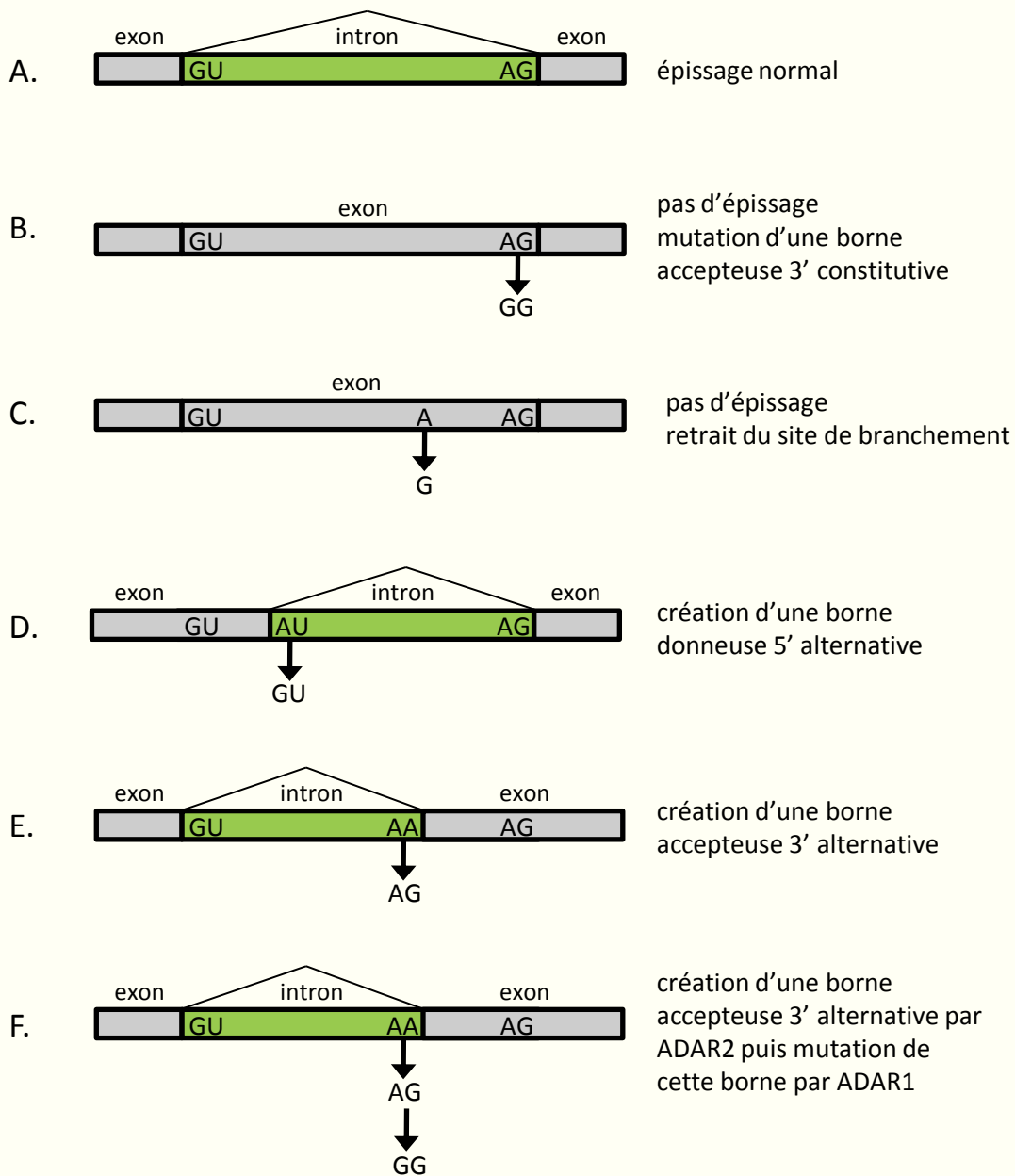


Figure 10 : Altération de l'épissage alternatif d'un transcrit par les protéines ADAR.

(A) Epissage normal du transcrit. (B) L'édition affecte le site 3' accepteur et abolit l'épissage du transcrit. (C) L'édition affecte le site de branchement et inhibe l'épissage du transcrit. (D) L'édition crée un site alternatif en 5'. (E) L'édition crée un site alternatif en 3'. (F) La protéine ADAR2 édite son propre transcrit créant un site accepteur 3' aboutissant à une protéine non fonctionnelle. Ce site peut ensuite être supprimé sous l'action d'ADAR1.

que les miARN, suggérant un rôle important de l'édition par les ADAR dans la régulation des processus post-transcriptionnels.

La plupart des sites d'édition sont préférentiellement édités par l'une ou l'autre des protéines ADAR1 et 2, laissant supposer une différence dans la reconnaissance de leurs substrats ARN. Cette différence pourrait s'expliquer par le nombre et l'espacement des motifs dsRBD des deux protéines (Figure 9), ainsi que par leur homodimérisation, nécessaire à leur activité catalytique (Cho *et al.*, 2003). En effet, il semble que la coopérativité des régions dsRBD au sein des homodimères serait responsable de la spécificité de ces protéines pour leur substrat (Valente & Nishikura, 2007). Enfin, aucune cible n'a été identifiée pour la protéine ADAR3, en raison de son absence d'activité catalytique *in vitro*. Il semblerait que son homodimérisation ne puisse avoir lieu que sous la dépendance de facteurs spécifiques du cerveau (Cho *et al.*, 2003).

Par contre, il est intéressant de noter que la spécificité des protéines ADAR1 et ADAR2 ne semble pas dépendre de cofacteurs spécifiques, car les profils d'édition obtenus *in vitro* sont identiques à ceux identifiés *in vivo* (Polson *et al.*, 1996). Ainsi, la plupart des longs ARNdb (>100pb) sont édités de façon multiple avec environ 50% des adénosines converties en inosines, tandis que les régions ARNdb courtes (20-30pb) ou les longs ARN possédant des mésappariements, des renflements ou des boucles, vont être édités spécifiquement avec seules quelques adénosines particulièrement ciblées (Nishikura, 2010). En effet, bien que les protéines ADAR semblent avoir une légère préférence pour des adénosines dans un contexte nucléotidique particulier, il apparaît que la présence de boucles ou de mésappariements semble essentielle afin de conférer une spécificité de reconnaissance de l'enzyme pour son substrat mais aussi pour spécifier les résidus à éditer (Dawson *et al.*, 2004; Lehmann & Bass, 1999).

1.2.3.2.2.3. Rôles des protéines ADAR

1.2.3.2.2.3.1. Rôle des ADAR dans la physiologie cellulaire

Bien que peu de cas d'édition dans les séquences codantes aient été recensés, les protéines ADAR jouent un rôle clef dans l'édition des transcrits codant pour des protéines du système nerveux central, comme de nombreux canaux ioniques et des récepteurs de neurotransmetteurs (Jepson & Reenan, 2008). Par exemple, ADAR1 est capable d'éditer le transcrit codant pour le récepteur de la sérotonine de type 2C, permettant de modifier trois codons via cinq sites d'édition et selon le profil d'édition d'exprimer plus de 24 isoformes différentes du récepteur (Nishikura, 2010). Le rôle majeur d'ADAR2 semble être le ciblage et la régulation de la sous unité B du récepteur GluR, car la mutation létale chez les souris invalidées pour ADAR2 peut être complétée par l'expression du transcrit GluR édité (Higuchi *et al.*, 2000). En effet, l'édition du site Q/R a des conséquences sur l'épissage

correct du transcrit de *GluR-B* nécessaire à l'assemblage convenable des quatre sous-unités de ce récepteur (Jepson & Reenan, 2008).

Cependant, chez les primates, la majorité des sites d'édition via les protéines ADAR se retrouverait dans les éléments répétés et abondants tels les transposons Alu et LINE (Levanon *et al.*, 2004). En effet, ces éléments qui sont retrouvés dans les 5' et 3' UTR forment grâce à leurs séquences répétées inversées, de grandes structures double brin en tige-boucle. Le rôle de l'édition de ces rétrotransposons est peu élucidé mais il pourrait être impliqué dans le développement et la différenciation de cellules embryonnaires (Osenberg *et al.*, 2010) et il pourrait également prévenir l'exonisation aberrante des séquences Alu au sein de l'ARNm mature (Sakurai *et al.*, 2010).

1.2.3.2.2.3.2. Rôle des ADAR dans la lutte antivirale

L'implication des protéines ADAR dans la lutte antivirale est due à leur spécificité d'action sur les ARNdb. En effet, les grands ARNdb sont rares dans une cellule et ils sont le plus souvent le signe d'une infection virale. Classiquement, la reconnaissance de ces ARNdb par des récepteurs tels que TLR3 (Toll Like Recepteur 3) mène à la réponse immunitaire innée et à la production d'interféron de type I (IFN α/β) et de type II (IFN γ). En effet, les interférons α et γ induisent la réponse primaire en stimulant la production de nombreuses protéines ciblant les ARNdb comme la PKR (RNA-dependent Kinase), la RNase L (2-5-A-dependent RNase), l'OAS (2'-5' oligoadenylate synthetase) ainsi que la protéine ADAR1-L.

A ce jour, seuls quelques virus et principalement des virus à ARN ont été décrits édités par la famille des ADAR : trois rétrovirus (VIH-1, Rous-associated virus (RAV-1) et Avian Leukosis Virus (ALV)) (Felder *et al.*, 1994; Hajjar & Linial, 1995; Phuphuakrat *et al.*, 2008), deux rhabdovirus (vesicular stomatitis virus et sigma virus) (Carpenter *et al.*, 2009; O'Hara *et al.*, 1984), deux paramyxovirus (Respiratory Syncytial Virus (RSV) et rougeole) (Rueda *et al.*, 1994; Toth *et al.*, 2009), le LCMV (Lymphocytic Choriomeningitidis Virus) (Zahn *et al.*, 2007) et le virus de l'hépatite delta (HDV) (Wong & Lazinski, 2002). Parmi les virus à ADN, seuls des transcrits des deux herpesvirus HHV4 (Iizasa *et al.*, 2010) et HHV8 (Gandy *et al.*, 2007) ou du polyomavirus SV40 (Kumar & Carmichael, 1997) ont été retrouvés édités. Il faut noter que dans le cas de HDV et des herpesvirus, qui utilisent l'édition pour la réalisation de leur cycle de réplication, la déamination A \rightarrow I présente une spécificité de site, alors que pour la plupart des autres virus les mutations retrouvées sont multiples et sembleraient correspondre à de l'hyper-édition dans le cadre d'une lutte anti-virale.

Par exemple, lors de la réplication de l'arénavirus LCMV, qui est un virus à ARN simple brin négatif segmenté, le fragment génomique S est édité par ADAR1-L lorsque le génome est sous une forme ARNdb transitoire. Cette hyper-édition pouvant toucher jusqu'à 38% des adénines des ARNm des glycoprotéines aboutit à la traduction de glycoprotéines non fonctionnelles ne pouvant atteindre la

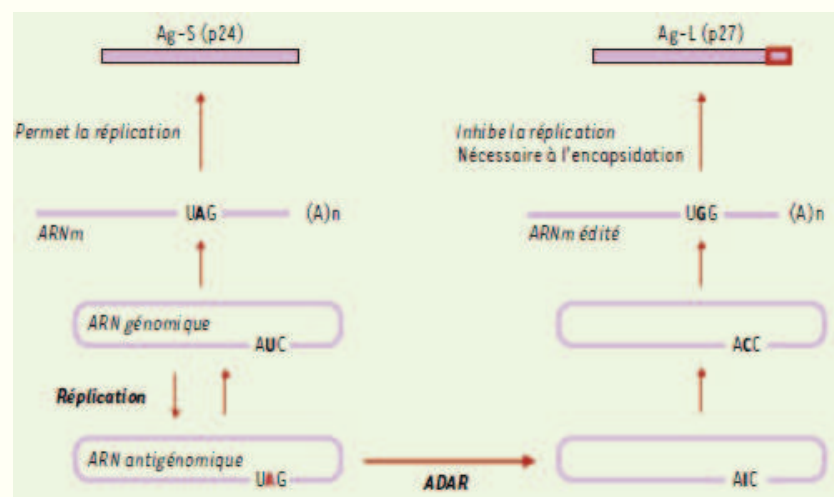


Figure 11 : Mécanisme d'édition de l'ARNm du virus de l'hépatite delta (Blanc et al., 2002). L'édition A=>I réalisée par ADAR1-S sur l'ARN anti-génomique permet la suppression d'un codon STOP. L'Ag-L (p27 ; 214 aa) est alors produit à la place de l'Ag-S (p24 ;195 aa).

lumière du réticulum endoplasmique ou la surface cellulaire, diminuant ainsi le pouvoir infectieux de la population virale (Zahn *et al.*, 2007).

Enfin, la défense antivirale induite par ADAR peut être détournée par les virus afin de servir leur propre cycle réplcatif. Par exemple, dans le cas du virus de l'hépatite delta (HDV) l'édition par ADAR1-S d'une partie des intermédiaires de réplication permet la poursuite du cycle réplcatif du virus (Wong & Lazinski, 2002). En effet, au cours du cycle réplcatif d'HDV, qui est un virus à ARNs_b défectif et satellite du virus de l'Hépatite B, deux antigènes Ag-S de 195 aa et Ag-L de 214 aa identiques dans leur partie NH₂ terminale sont produits. En absence d'édition, la traduction du messenger conduit à l'expression de l'Ag-S qui participe au déclenchement de la réplication virale. Cependant, une partie des intermédiaires de réplication subissent une désamination supprimant un codon stop sur les messagers correspondant à ces génomes modifiés entraînant la traduction de l'Ag-L (Figure 11). La synthèse de ce dernier inhibant la réplication et permettant l'encapsidation donne à l'édition un rôle décisif dans le déroulement du cycle viral (Blanc V, 2002).

Chez le virus HHV8, l'édition A→G induite par ADAR1 élimine le pouvoir transformant induit par le transcrit codant la kaposine A et contenant également KSHV-miR-K12-10 (Gandy *et al.*, 2007). L'édition de ce nucléotide serait capable d'induire un changement d'un aa de la kaposine A et C et de modifier le second nucléotide du miARN KSHV-miR-K12-10. De plus, le taux d'édition semble être régulé par le virus et il serait plus important lors des stades de réplication du virus comparativement aux phases de latence. En effet, dix fois plus de transcrits édités sont retrouvés lors de l'induction de la réplication lytique (Gandy *et al.*, 2007).

1.2.4. Le mécanisme d'épissage

Chez l'ensemble des eucaryotes, ainsi que chez les archéobactéries et les cyanobactéries, les gènes sont organisés en une suite d'introns et d'exons alternés, dont généralement, les introns représentent les parties non codantes et les exons les parties codantes. L'épissage (splicing) consiste en la maturation des ARNm durant laquelle les introns sont éliminés, tandis que les exons sont conservés afin de donner un ARNm dit mature ultérieurement traduit en protéine. Ce phénomène, assuré par un complexe ribonucléoprotéique appelé spliceosome, se déroule dans le noyau de la cellule en concomitance avec la transcription.

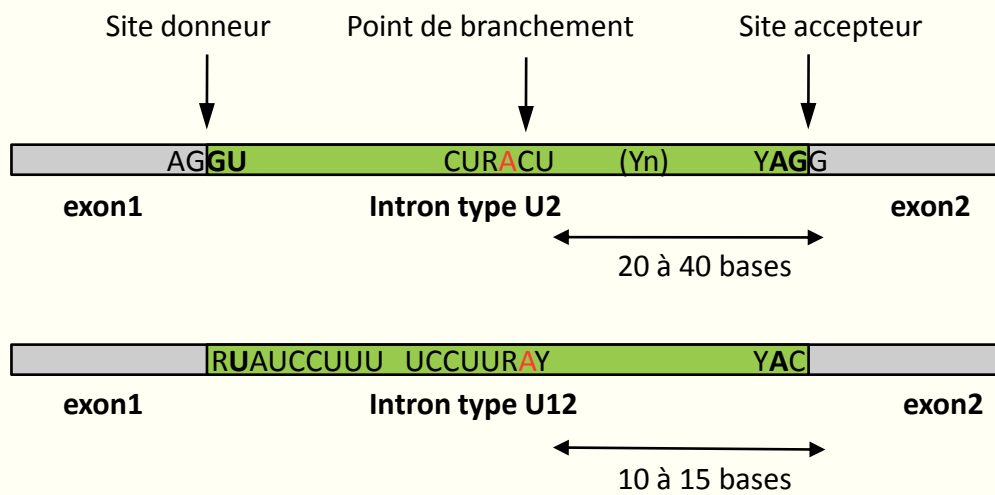


Figure 12 : Comparaison des séquences régulatrices consensus de type U2 et U12. L'adénosine du point de branchement est indiquée en rouge et les nucléotides très conservés sont indiqués en caractère gras.

1.2.4.1. Le splicéosome

1.2.4.1.1. Les séquences consensus d'épissage

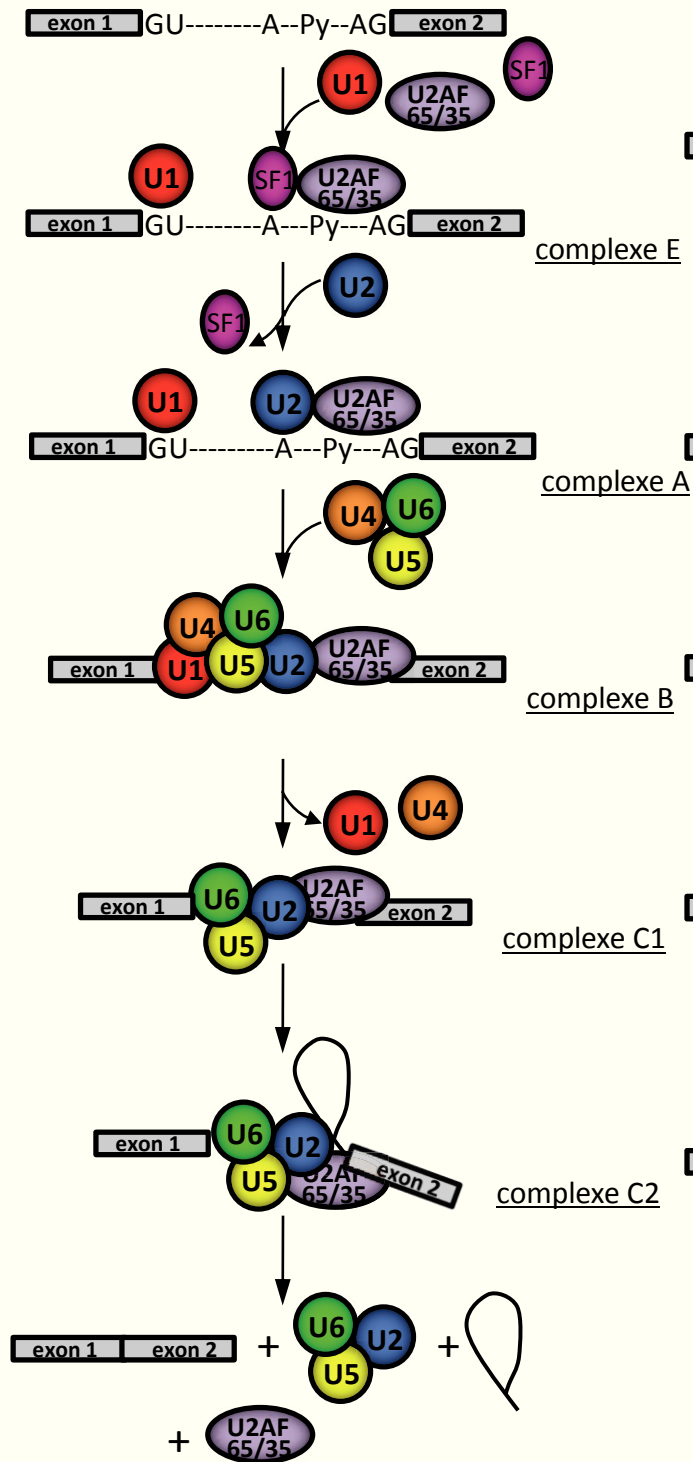
Le splicéosome est capable de reconnaître trois séquences régulatrices indispensables au recrutement de la machinerie d'épissage, localisées au niveau des jonctions intron/exon et à l'intérieur de l'intron. La jonction exon/intron en 5' de l'intron est appelée site donneur (5' Splice Site), la jonction intron/exon en 3' de l'intron est appelée site accepteur (3' Splice Site), et la séquence interne de l'intron, permettant la formation du lariat, est appelée site de branchement. Ce dernier correspond à une adénine localisée 10 à 40 bases environ en amont de l'extrémité 3' de l'intron suivie, dans le cas des introns de type U2, d'une région riche en pyrimidines de composition et de longueur variables.

En effet, il existe deux types de splicéosome différents au sein des eucaryotes. Le splicéosome majoritaire dit de type U2 et le splicéosome minoritaire de type U12 qui est retrouvé pour 1% des gènes humains (Levine & Durbin, 2001). Contrairement au splicéosome de type U2 qui est retrouvé chez l'ensemble des eucaryotes, le splicéosome de type U12 n'est présent que chez les vertébrés, certains champignons, nématodes et plantes (Bartschat & Samuelsson, 2010).

Les introns de type U2 présentent des caractéristiques très conservées. Ainsi, le site donneur correspond presque systématiquement à une séquence AG/GU, puisque le deuxième site le plus rencontré (AG/GC) représente moins de 0,5% des sites donneurs référencés à ce jour (Burset *et al.*, 2001). Enfin, la séquence consensus du site accepteur est toujours de type YAG/R et l'adénine du site de branchement est comprise dans la séquence CURACU (Figure 12).

Le second type d'intron, de type U12 (Hall & Padgett, 1994; Jackson, 1991), se caractérise par une séquence /RUAUCCUUU au niveau du site donneur et une séquence de branchement de type UUCCUURAY (Figure 12) (Alioto, 2007). A l'instar des introns de type U2, les bornes des introns de type U12 les plus représentées sont de type « GU-AG ». Cependant les bornes de type « AU-AC » représentent ici plus d'un tiers des bornes de type U12 référencées. De plus, bien que retrouvées à faible taux, les différentes bornes fonctionnelles de type « AU-AA ; AU-AG ; AU-AU ; GU-AA ; GU-AU et GU-GG » ont également été décrites (Dietrich *et al.*, 2005; Levine & Durbin, 2001). Enfin, il faut noter l'existence d'un fort taux d'erreur au niveau du site accepteur lors des épissages de type U12 avec des décalages pouvant atteindre jusqu'à 6 nucléotides en aval pour le début de l'exon (Levine & Durbin, 2001).

Splicéosome de type U2



Splicéosome de type U12

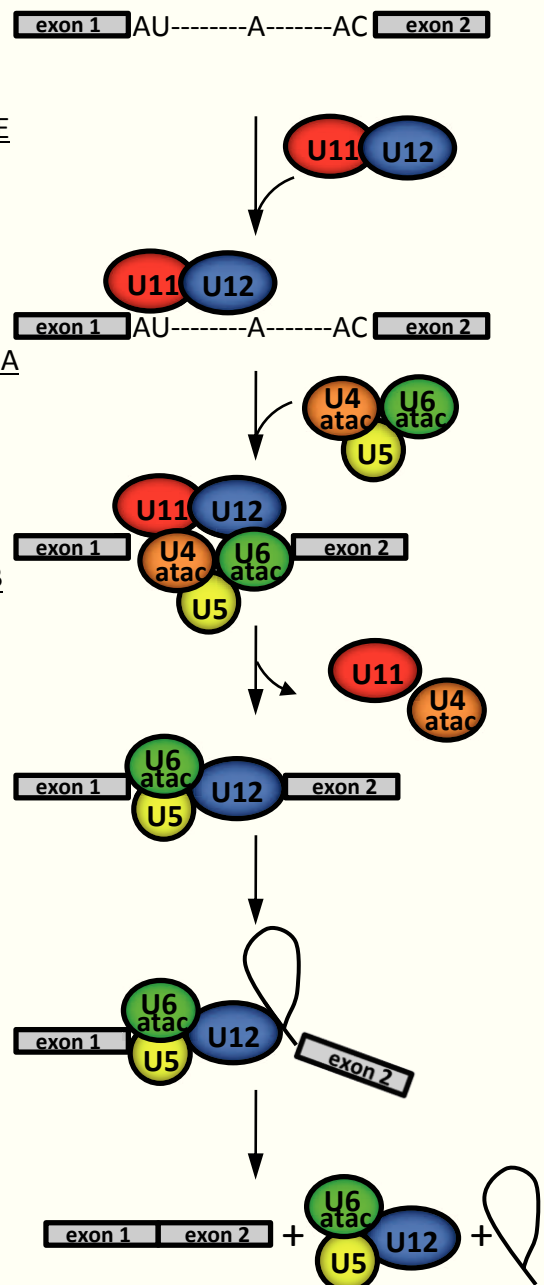


Figure 13 : Comparaison des différentes étapes de l'assemblage des splicéosomes de type U2 et U12.

1.2.4.1.2. Composition des spliceosomes

D'une manière générale, le spliceosome est un complexe dynamique dans lequel plus de 300 protéines différentes ont été identifiées (Jurica & Moore, 2003). Les deux types de spliceosome comportent 5 snRNP (small nuclear ribonucleoproteins). Cependant, seule la snRNP U5 est présente au sein des deux spliceosomes, les 4 autres snRNP rencontrées pour le spliceosome de type U2 étant les snRNP U1, U2, U4 et U6, et les snRNP U11, U12, U4atac et U6atac pour le spliceosome de type U12. Par ailleurs, deux autres facteurs, l'hétérodimère U2AF (U2AF65 et U2AF35, U2 snRNP Auxiliary Factor) et le facteur SF1 (Splicing factor 1) principalement impliqués dans le recrutement des snRNP sont essentiels pour le spliceosome de type U2.

Chaque snRNP renferme un petit ARN nucléaire spécifique riche en uridine (UsnARN : Uridine rich small RNA). De taille variable (100 à 200 nt chez l'homme), ces UsnARN subissent de fortes modifications post transcriptionnelles. Ils sont partiellement complémentaires de la région bordant le site d'épissage et surtout ils portent l'activité catalytique du snRNP (Will & Luhrmann, 2001). Après être généralement transcrits par l'ARN Pol II, les UsnARN sont exportés dans le cytoplasme où ils vont s'associer à sept protéines Sm ou avec des protéines apparentées, qui formeront un anneau heptamérique et ainsi le corps de la snRNP (Will & Luhrmann, 2001). Dans le cytoplasme, les UsnARN subissent une hyper-méthylation de leur coiffe permettant le retour des pré-snRNP dans le noyau. Ensuite, les snRNP transitent par les corps de Cajal où ils finissent leur maturation (méthylations et pseudouridylation), et où ils s'assemblent en di ou tri-snRNP avant de rejoindre les « speckles » et/ou les sites de transcription (Darzacq *et al.*, 2002; Jady *et al.*, 2003). Pour finir, deux familles de protéines régulatrices, les protéines SR (riches en dipeptides sérine/arginine) et la famille des hnRNP (Heterogeneous Nuclear RNP) sont importantes pour le déroulement des événements d'épissage. En règle générale, les protéines SR semblent agir en tant qu'activatrices alors que les hnRNP ont plutôt une action inhibitrice de l'épissage.

1.2.4.1.3. Mécanismes d'action des spliceosomes

Les deux types de spliceosome coexistent au sein des cellules eucaryotes et bien que leur composition en snRNP diffère, les mécanismes d'épissage sont très similaires (Schneider *et al.*, 2002; Turunen *et al.*, 2008).

L'épissage des introns se déroule en deux étapes de trans-estérification successives. Durant la première étape, le 2'OH de l'adénine du point de branchement est activé pour attaquer la liaison phosphodiester du site 5' donneur, aboutissant ainsi à la libération de l'exon 5' et à la formation de l'intron branché (le lasso ou lariat) contenant une liaison phosphodiester 2'-5'. La seconde étape

correspond à l'attaque nucléophile du 3'OH de l'exon 5' au niveau de la liaison phosphodiester du 3' accepteur, menant à la ligation des deux exons et à la libération de l'intron en lasso (Figure 13).

Deux modèles d'assemblage des spliceosomes sur l'ARN pré-messager ont été décrits. Un premier, plus documenté, montre un assemblage séquentiel des différents composants sur l'intron à épisser (Figure 13). Ainsi, l'épissage induit par le spliceosome de type U2 débute par l'interaction du snRNP U1 avec le site donneur 5', du facteur SF1 avec le point de branchement et du dimère U2AF avec la région riche en pyrimidine (U2AF65) et avec le site accepteur 3' (U2AF35) pour former le complexe E. Le remplacement du facteur SF1 par le snRNP U2 au niveau de l'adénine du point de branchement aboutit ensuite à la formation du complexe A, puis au complexe B par adjonction de la tri-snRNP U4/U5/U6. Après de multiples réarrangements ARN-ARN et ARN-protéine, l'interaction entre les snRNP U4 et U6 est rompue permettant à snRNP U6 de s'apparier avec le site donneur en déplaçant la snRNP U1. Après ces réarrangements provoquant la déstabilisation et la dissociation des snRNP U1 et U4 du spliceosome, le complexe catalytique est alors formé des snRNP U2/U5/U6. La première réaction de trans-estérification a lieu suite au rapprochement du site accepteur et du site de branchement consécutif à l'interaction entre U2 et U6 (complexe C). Le positionnement des exons pour la seconde réaction de trans-estérification conduisant à la ligation des exons et l'excision des introns (Newman, 1997) est ensuite réalisé par U5. Finalement, le spliceosome se dissocie, relâchant les snRNP qui seront recyclées pour un nouveau processus d'épissage. Par ailleurs, la participation des protéines SR s'effectue essentiellement lors de la formation des complexes B et C, en favorisant l'incorporation du tri-snRNP U4/U5/U6 (Rosciigno & Garcia-Blanco, 1995) puis l'appariement entre les bases nucléotidiques de U2 et U6 (Tarn & Steitz, 1995).

A l'exception de la première étape, l'assemblage du spliceosome de type U12 est similaire à celui de type U2. En effet, avant de s'associer au pré-ARNm, les snRNP U11 et U12 forment un complexe très stable qui se lie au site accepteur et au site de branchement conduisant directement à la formation du complexe A (Frilander & Steitz, 1999; Turunen *et al.*, 2008). Ensuite, les snRNP du spliceosome de type U12, qui possèdent de grandes homologues de structure et de fonction avec celles du spliceosome de type U2 vont interagir à l'identique de ces dernières (Figure 13) (Schneider *et al.*, 2002; Shukla & Padgett, 2001).

Le second modèle consiste en l'existence d'un spliceosome pré-assemblé, pouvant se fixer directement sur le pré-ARNm. En effet, un grand complexe a été retrouvé chez la levure dénommé holospliceosome, composé des cinq snRNP U1, U2, U4, U5 et U6 et d'une soixantaine de protéines associées en absence de pré-ARNm (Stevens *et al.*, 2002). Ainsi, les différents complexes spliceosomaux intermédiaires E, A, B et C correspondraient à des étapes de stabilisation d'interactions plutôt qu'au recrutement séquentiel des facteurs d'épissage (Rino & Carmo-Fonseca, 2009).

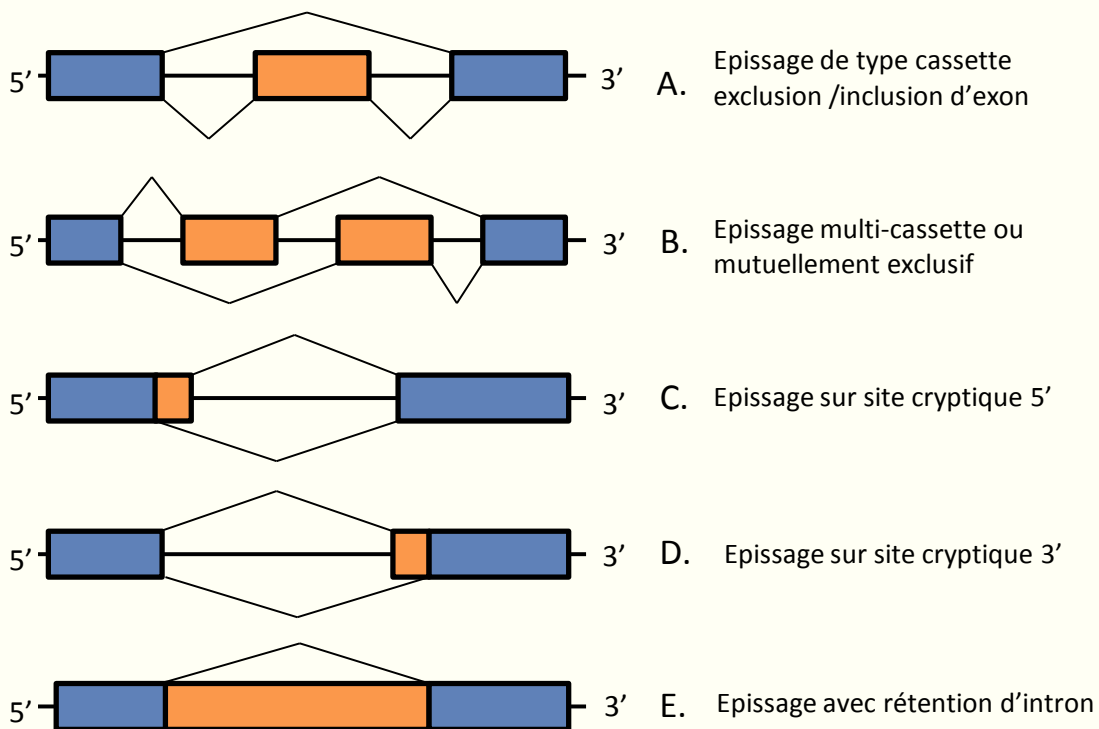


Figure 14 : Représentation schématique des différents types d'épissage alternatif. Les exons constitutifs sont représentés en bleu et les exons alternatifs en orange.

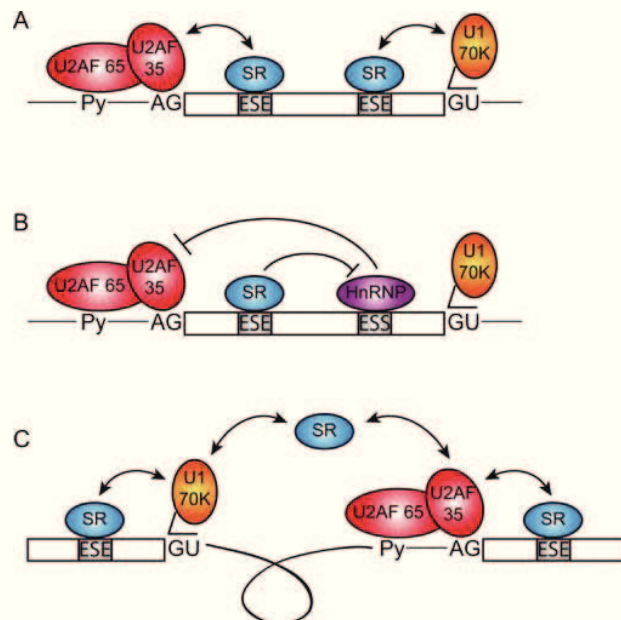


Figure 15 : Rôles des protéines SR dans l'épissage. (D'après Long et al., 2009)
 (A) Recrutement des facteurs U2AF35 et U1 par la fixation d'une protéine SR sur un site ESE. (B) Inhibition de l'effet répresseur d'un facteur hnRNP fixé au niveau d'un ESS. (C) Création d'un réseau d'interaction par les protéines SR aidant à la juxtaposition des sites d'épissage par interaction avec U1 et U2AF35.

1.2.4.2. L'épissage alternatif

L'épissage alternatif est un processus naturel permettant la génération de plusieurs ARNm matures différents, codant ou non codant pour des isoformes protéiques à partir d'un même pré-ARNm. Il est considéré comme la source la plus importante de diversité des protéines chez les vertébrés (Graveley, 2001) et sa régulation, qui est sous le contrôle de très nombreux facteurs, est essentielle à la cellule.

Cinq grands types d'épissage alternatifs peuvent être rencontrés (Figure 14). L'épissage du type cassette correspond à l'inclusion ou à l'exclusion de l'ARNm, d'un exon alternatif compris entre deux exons constitutifs (Figure 14A). En effet, il existe deux types d'exons, les exons dits constitutifs qui seront toujours inclus dans l'ARNm et les exons dits alternatifs qui pourront se retrouver épissés avec les introns. Dans le cas où deux exons alternatifs exclusifs l'un de l'autre sont compris entre deux exons constitutifs, la machinerie d'épissage peut déterminer lequel des deux exons sera épissé aboutissant à un épissage dit multi-cassette ou mutuellement exclusif (Figure 14B). L'épissage via des sites cryptiques en 5' ou en 3' permet la délétion d'une partie d'un exon ou la conservation d'une partie de l'intron (Figure 14C et D). Enfin, il peut y avoir absence d'épissage, ce qui provoque la rétention de l'intron dans l'ARNm (Figure 14E). Comme l'ensemble de ces schémas d'épissage peuvent potentiellement être présents au sein d'un gène donné, de très nombreuses combinaisons d'ARNm potentiels peuvent être générées à partir de ce gène unique.

1.2.4.2.1. Les séquences régulatrices

En plus des éléments précédemment décrits nécessaires à sa mise en place, la formation du spliceosome est sous le contrôle d'éléments *cis* correspondant à des séquences régulatrices activatrices ou inhibitrices. Ces séquences peuvent être localisées aussi bien dans les exons que dans les introns. Ainsi quatre catégories d'éléments *cis* sont retrouvées : les ESE et ESS (Exonic Splicing Enhancer et Exonic Splicing Silencer) et les ISE et ISS (Intronic Splicing Enhancer et Intronic Splicing Silencer). Ces séquences auxiliaires présentent des caractéristiques communes : elles sont courtes (moins de dix nucléotides), variables en terme de séquence, souvent présentes en plusieurs copies et faiblement actives individuellement. La plupart d'entre elles sont conservées entre les espèces, cependant elles contiennent des motifs dégénérés ce qui les rend difficiles à identifier par les approches de bioinformatique (Ladd & Cooper, 2002).

Les ESE sont reconnus principalement par les protéines SR qui permettent le recrutement du spliceosome (de type U2 ou U12) au niveau de l'intron adjacent (Figure 15) (Hastings & Krainer, 2001; Long & Caceres, 2009). Une fois fixées sur la séquence ESE, les protéines SR permettraient le recrutement et la stabilisation des interactions entre U1 et le site donneur, ou entre U2AF35 et le site

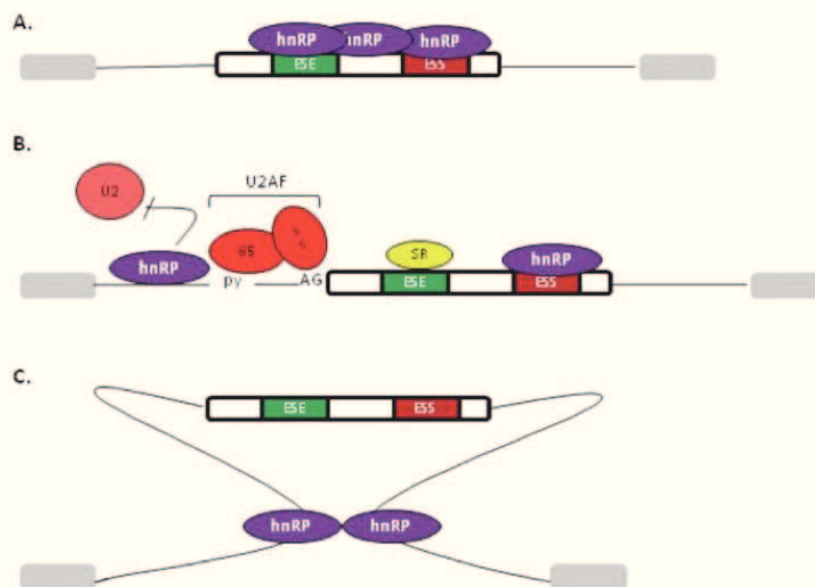


Figure 16 : Rôles de la protéine hnRNP A1 dans l'épissage.

(A) Blocage de l'accès à un site ESE par agglomération de la hnRNP A1. (B) Inhibition du recrutement de U2 par hnRNP A1. (C) Création d'une boucle entre deux introns par interaction entre deux hnRNP A1, permettant l'exclusion d'un exon.

accepteur et entre le snARN du complexe U2 et le site de branchement (Figure 15A) (Long & Caceres, 2009). De plus, elles sembleraient pouvoir bloquer l'action d'un facteur inhibiteur de l'interaction entre U2AF et le site accepteur (Figure 15B) (Zhu *et al.*, 2001). Par ailleurs, en interagissant avec U1 et U2AF35, les protéines SR peuvent également former un réseau d'interactions protéine-protéine visant à juxtaposer les sites donneurs et accepteurs durant l'assemblage du spliceosome (Figure 15C) (Long & Caceres, 2009).

Les ISE sont beaucoup moins documentés et leur position au sein de l'intron ainsi que leur séquence sont relativement variables. Les ISE peuvent activer fortement l'épissage mais avec un mécanisme probablement différent de celui des ESE (Black, 2003). Les protéines reconnaissant ces motifs sont également peu caractérisées. Certaines semblent appartenir à la famille des protéines SR et d'autres seraient des membres de la famille CELF (CUGBP and ETR-Like Factors) (Black, 2003).

Les ESS sont généralement reconnus par des membres de la famille des hnRNP, tandis que les ISS fixent principalement une protéine interagissant avec la région poly-pyrimidique (Black, 2003). Le membre le plus étudié de la famille des hnRNP, hnRNP A1 est capable de réprimer l'épissage selon différents mécanismes : a) en se multimérisant le long des exons et en empêchant ainsi l'accès aux ESE (Figure 16A) (Zhu *et al.*, 2001), b) en bloquant le recrutement des snRNP (Figure 16B) (Tange *et al.*, 2001), c) en créant une boucle dans un intron pour rapprocher deux exons adjacents et provoquer l'inclusion d'un exon à l'intérieur de cette boucle (Figure 16C) (Martinez-Contreras *et al.*, 2006).

1.2.4.2.2. L'architecture des exons et des introns

L'architecture des exons et des introns influence également la reconnaissance des sites d'épissage, la formation du spliceosome et la fréquence de l'épissage alternatif (Hertel, 2008). D'une manière générale, la reconnaissance des sites d'épissage est plus efficace si les exons et les introns sont petits. Une augmentation de la taille de l'exon induit une augmentation de l'épissage alternatif au sein de cet exon. Cependant, il semble que les grands exons seront gardés si les introns les entourant sont petits. De plus, un exon quelle que soit sa taille, entouré de grands introns subira généralement un épissage alternatif plus important que s'il est entouré de petits introns.

Devant l'hétérogénéité de taille des introns et des exons, il a été supposé que le spliceosome puisse reconnaître les sites d'épissage selon deux mécanismes appelés la définition de l'exon et la définition de l'intron (Berget, 1995). On parle de définition de l'intron lorsque les bornes 5' et 3' de celui-ci sont directement reconnues par la machinerie d'épissage et de définition de l'exon quand se sont les bornes de l'exon qui sont reconnues. Des études *in-vitro* ont montré que l'intron semble être défini si sa taille est comprise entre 200 et 250 nucléotides et qu'en dehors de cette taille, ce serait plutôt l'exon qui serait reconnu (Fox-Walsh *et al.*, 2005). Dans ce dernier cas, la machinerie d'épissage localise une paire de sites d'épissage 5' et 3' rapprochée pouvant appartenir à un même

exon. Ainsi, quand l'exon est défini, les exons voisins se rapprochent les uns des autres, probablement par l'intermédiaire d'interactions entre les facteurs reconnaissant les exons, puis l'épissage se déroule classiquement.

1.2.4.2.3. Les structures locales de l'ARNm

Les structures secondaires et tertiaires de l'ARNm peuvent également jouer un rôle significatif dans l'assemblage du spliceosome (Hertel, 2008). En effet, ces repliements de l'ARNm, impliquant jusqu'à plusieurs centaines de nucléotides, peuvent influencer la reconnaissance des sites d'épissage et des sites activateurs ou inhibiteurs en les masquant (Hertel, 2008).

1.2.4.2.4. Régulation de l'épissage par l'ARN Pol II

Il existe aussi un modèle de régulation de l'épissage impliquant le complexe d'initiation de la transcription et l'ARN Pol II (Hertel, 2008). Du fait que l'épissage soit un mécanisme co-transcriptionnel, ce modèle suggère que les facteurs de l'épissage sont associés à la transcription par l'ARN Pol II et qu'ainsi, des différences au niveau de la structure promotrice pourraient induire un recrutement de facteurs de l'épissage spécifiques (Hertel, 2008). De plus, l'ARN Pol II fonctionnant dans le sens 5' vers 3', des exons alternatifs peuvent alors être transcrits avant ou après la synthèse d'exons voisins compétiteurs. Ainsi, la vitesse relative de production d'un exon compétiteur peut induire des modifications du profil d'épissage (Hicks *et al.*, 2006).

1.2.4.2.5. Les sites d'épissages cryptiques

Des séquences consensus correspondant à des sites s'épissage peuvent être retrouvées à l'intérieur d'un exon ou d'un intron. Ces sites d'épissage, qui ne sont pas les sites classiques, sont appelés des sites d'épissage cryptiques et leur sélection par le spliceosome semble être sous la dépendance des séquences régulatrices ESE, ESS, ISE et ISS. L'utilisation d'un site cryptique entraîne la production d'un ARNm mature pouvant contenir une partie de l'intron adjacent ou ayant perdu une partie de l'exon. Ce changement induit l'apparition de nouvelles isoformes de la protéine si le cadre de lecture est maintenu ouvert ou d'un ARNm non fonctionnel suite à l'apparition d'un codon d'arrêt de traduction prématuré. A titre d'exemple, des profils spécifiques d'utilisation des sites cryptiques ont été observés *in vivo* durant la différenciation des érythrocytes (Haj Khelil *et al.*, 2008).

1.2.4.3. Devenir des ARNm épissés reconnus par la voie du nonsense mediated decay (NMD)

Environ 80% des événements d'épissage alternatifs changeraient la nature de la protéine, mais seulement 20% conduiraient à la production de protéines tronquées suite à un changement du cadre de lecture et à l'apparition d'un codon stop prématuré (Modrek & Lee, 2002). Ainsi, dans la plupart des cas, ces ARNm tronqués sont pris en charge par le mécanisme de NMD (Nonsense Mediated mRNA Decay). D'une manière générale, ce mécanisme prend en charge les ARNm possédant des aberrations dans la position du codon stop (Baker & Parker, 2004). Dans la majorité des cellules de mammifères, le NMD est déclenché si un codon stop est suivi d'une jonction exon-exon située à plus de 50-55 nucléotides en aval (Nagy & Maquat, 1998).

2. Régulation de l'expression génique par les microARN

2.1. Généralités

Les microARN (miARN) sont une classe de petits ARN non codant simple brin d'environ 22 nucléotides, jouant un rôle clef dans la régulation de l'expression génique au niveau post-transcriptionnel par un mécanisme d'ARN interférence. Les miARN ont été découverts chez le nématode *Caenorhabditis elegans* en 1993 (Lee *et al.*, 1993), mais ce n'est qu'en 1998 que fut identifié le phénomène d'ARN interférence (Fire *et al.*, 1998; Tabara *et al.*, 1998; Timmons & Fire, 1998) ; puis durant ces dix dernières années, de nombreux miARN ont été identifiés dans les mondes animal, végétal et viral où ils ont été reconnus comme des éléments régulateurs essentiels de nombreuses fonctions cellulaires incluant la différenciation, le développement, la prolifération ou encore l'apoptose.

Le mécanisme d'interférence entraîne la diminution de l'expression d'un gène par un petit ARN interférant (ARNi) possédant une séquence entièrement ou partiellement complémentaire de l'ARNm de ce gène. Ces ARNi dont font partie les miARN peuvent agir à plusieurs niveaux sur l'ARNm : soit en bloquant sa traduction ou en accélérant sa dégradation par déadénylation lors d'une complémentarité imparfaite, soit en le clivant au niveau de l'interaction miARN/ARNm lors d'une complémentarité parfaite. Cependant, certains cas d'activation de la transcription (Place *et al.*, 2008) ou de la traduction (Henke *et al.*, 2008; Orom *et al.*, 2008; Vasudevan *et al.*, 2007; 2008) par les miARN ont également été recensés montrant l'étendue du pouvoir d'action de ces petits ARN.

Chez l'homme, plus de 800 miARN ont été identifiés et plus de 60% des gènes codant des protéines chez les mammifères semblent être régulés par les miARN (Fabian *et al.*, 2010). Ils sont exclusivement retrouvés chez les eucaryotes, cependant des homologues des protéines Argonaute

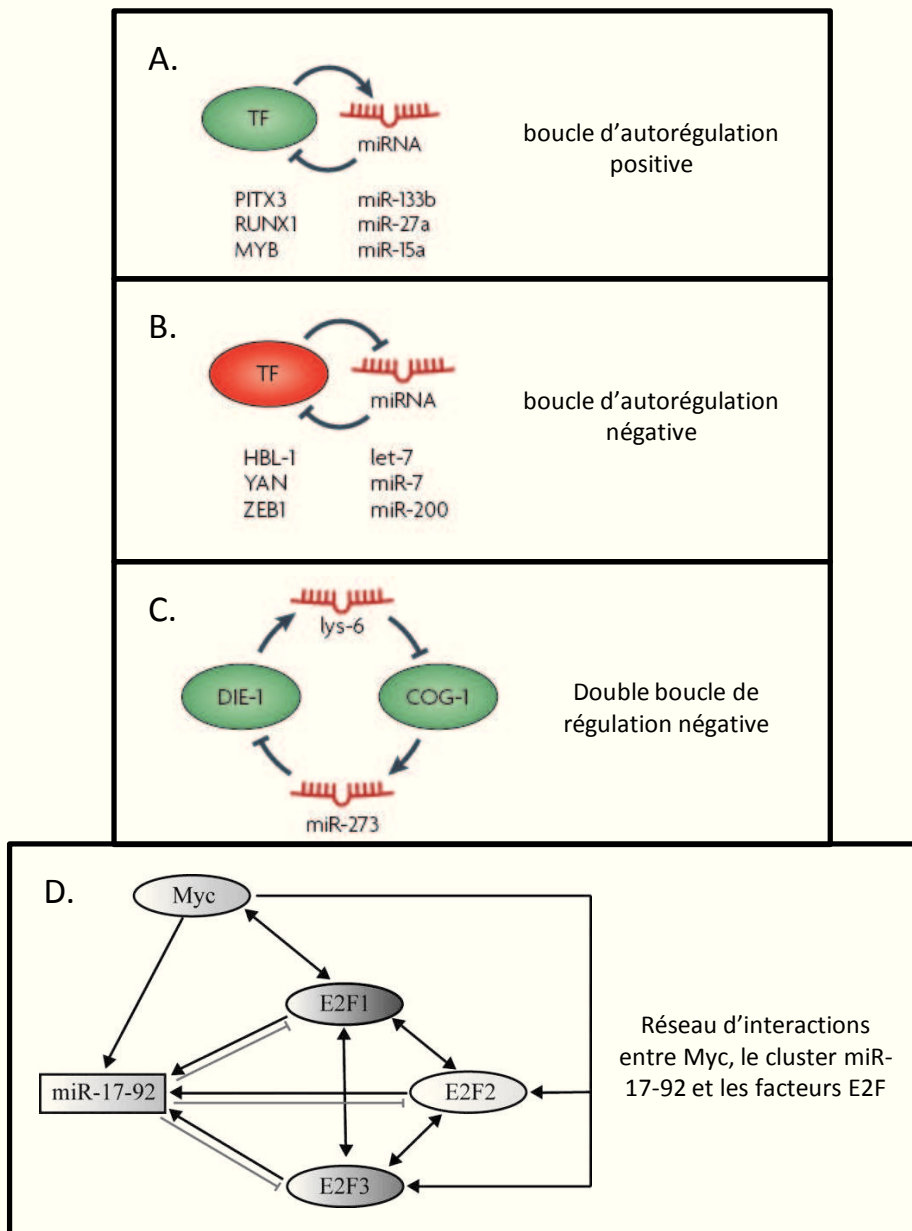


Figure 17 : Représentation schématique de différentes boucles de régulation entre les microARN et les facteurs de transcription (d'après Krol et al 2010).

(A) Représentation d'une boucle d'autorégulation positive retrouvée entre les miR-133b, miR-27a, miR-15a et respectivement PITX3, RUNX1, MYB (B) Représentation d'une boucle d'autorégulation négative identifiée entre les miARN let-7, miR-7, miR-200 et respectivement HBL-1 YAN, ZEB1. (C) Représentation d'une double boucle de régulation négative retrouvée chez *Caenorhabditis elegans*, entre les miARN miR-273 et lys-6 et des facteurs COG-1 et DIE-1. (D) Interactions entre c-Myc, les facteurs E2F et le cluster de miARN 17-92 (Xiang et al., 2010). Les flèches bidirectionnelles représentent une induction mutuelle de la transcription, et les flèches à tête plates font référence à une inhibition transcriptionnelle ou à une répression induite par le cluster 17-92.

(AGO) qui jouent un rôle primordial dans le phénomène d'ARN interférence ont également été identifiées chez les bactéries et les archaeobactéries (Carthew & Sontheimer, 2009).

2.2. Biogenèse des microARN

2.2.1. Transcription des microARN

Les gènes permettant l'expression des miARN sont principalement transcrits par l'ARN polymérase II (Lee *et al.*, 2004). Chez les vertébrés, la moitié des miARN sont transcrits sous forme intronique au sein de gènes codant des protéines (Davis & Hata, 2009), et parmi ces miARN introniques, environ 35% possèdent leurs propres promoteurs (Monteys *et al.*, 2010). Bien que minoritaire, la transcription des miARN par l'ARN polymérase III a également été mise en évidence notamment chez l'homme (Borchert *et al.*, 2006) ou certains virus (Pfeffer *et al.*, 2005; Reese *et al.*, 2010).

Les transcrits primaires de miARN (pri-miARN) possédant leurs propres promoteurs présentent des séquences régulatrices très similaires à celles codant des protéines (Corcoran *et al.*, 2009). Ils utilisent la même machinerie transcriptionnelle, possèdent les mêmes caractéristiques et sont régulés par les mêmes facteurs de transcription que l'ensemble des ARNm. De plus, à l'instar des ARNm, les pri-miARN sont capés, polyadénylés et fréquemment épissés ; cependant ils se différencient par la présence d'un ou plusieurs précurseurs de miARN (pré-miARN) caractérisés par la présence de structures tige-boucle d'une longueur d'environ 70 nucléotides dans laquelle sont généralement retrouvés deux miARN matures.

Il est intéressant de noter que les miARN participent régulièrement aux boucles d'autorégulation (feedback loop) positives ou négatives avec de nombreux facteurs de transcription (Figure 17) régulant leur biogenèse ou celle de leurs cibles (Krol *et al.*, 2010). Dans l'exemple du cluster miR-17-92 (Figure 17D), les miARN de ce cluster rentrent dans une boucle de régulation induite par c-Myc dans laquelle ce dernier active la transcription des facteurs de la famille E2F, tout en contrôlant leur traduction via les miARN du cluster-17-92. La régulation fine de ces facteurs conduit ainsi à l'inhibition du rôle proapoptotique des E2F et à la progression du cycle cellulaire (O'Donnell *et al.*, 2005; Xiang & Wu, 2010).

2.2.2. Maturation des microARN

L'obtention d'un miARN mature à partir du pri-miARN se fait en plusieurs étapes (Figure 18), mettant en jeu deux endoribonucléases III, une nucléaire Drosha et une cytoplasmique Dicer, ainsi

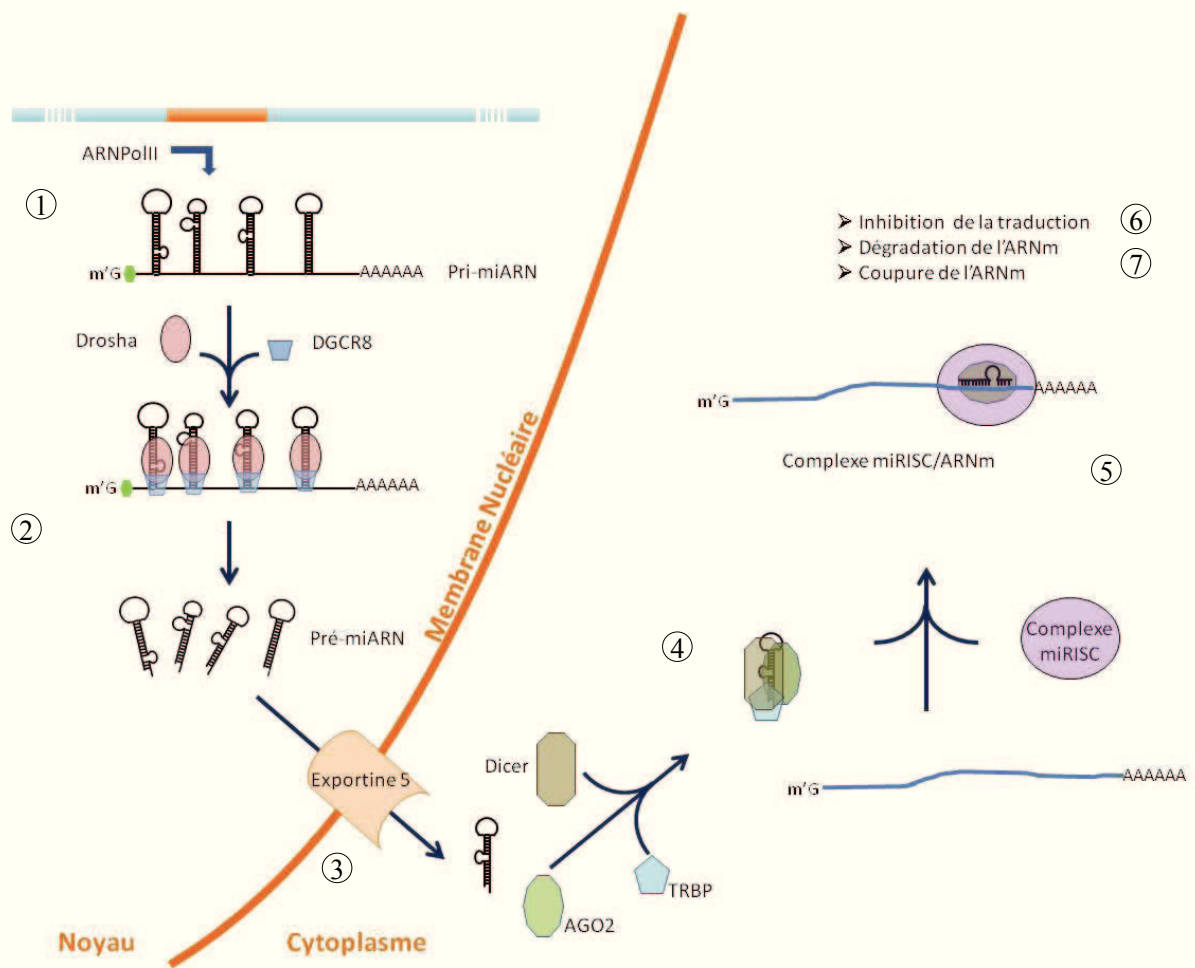


Figure 18 : Biogenèse des microARN.

Le gène codant le pri-miARN est transcrit par l'ARN polymérase II (1). Le complexe Droscha clive ce pri-miARN pour libérer un second précurseur, le pré-miARN (2) qui est exporté du noyau au cytoplasme via l'exportine 5 (3). Le complexe Dicer/TRBP/AGO clive la structure tige boucle de ce précurseur pour libérer un duplex de microARN (4). Un seul brin du duplex d'ARN est incorporé dans le complexe miRISC (5). Une fois chargé de ce brin, le complexe miRISC induit l'effet de répression post-transcriptionnelle soit par clivage des ARNm (6), soit par inhibition de la traduction et ou dégradation de l'ARNm (7).

qu'une protéine essentielle au transport cytoplasmique, l'Exportine-5. Ce n'est qu'après son incorporation au sein du complexe miRISC (microRNA-Induced Silencing Complex) que le miARN mature possèdera son pouvoir interférant.

2.2.2.1. Maturation du pri-microARN

La structure tige boucle du pri-miARN est primordiale pour l'entrée dans les différents processus de maturation des miARN. Cette structure, qu'elle soit isolée ou en cluster, issue d'un promoteur indépendant ou d'un intron (Morlando *et al.*, 2008), est reconnue et clivée co-transcriptionnellement par le même complexe ribonucléoprotéique, le microprocesseur.

Le microprocesseur correspond en l'association de DGCR8 (DiGeorge Syndrome Critical Region gene 8) et de Drosha (Figure 18), qui sont toutes deux nécessaires et suffisantes à la maturation des pri-miARN en pré-miARN (Gregory *et al.*, 2004). Le cofacteur DGCR8 permet la reconnaissance, la fixation et la bonne conformation du pri-miRNA (Yeom *et al.*, 2006). Cependant, d'autres protéines s'ajoutent à Drosha et DGCR8 afin de moduler la maturation des pri-miARN, telles que les hélicases p68 (Dead box RNA helicase DDX5) et p72 (DDX17) ainsi que les facteurs nucléaires NF90 et NF45 (Gregory *et al.*, 2004; Slezak-Prochazka *et al.*, 2010). Plus précisément, les facteurs NF90/NF45 semblent agir négativement sur cette maturation (Sakamoto *et al.*, 2009) tandis que les hélicases DDX5 et DDX17 semblent augmenter la maturation des pri-miARN en pré-miARN par Drosha (Fukuda *et al.*, 2007). De plus, des études ont montré que des interactions entre ces hélicases et d'autres protéines influencent également la maturation de différents pré-miARN. Par exemple, des études ont montré une augmentation de la maturation du pri-miR-21 par l'interaction entre certaines protéines SMAD et DDX5 (Davis *et al.*, 2008), ou encore une diminution de la maturation des miR-16, miR-125a, miR-143, miR-145 et miR-195 par l'interaction entre l'ER α (Estrogen Receptor α) et DDX5/DDX17 (Yamagata *et al.*, 2009).

Il est intéressant de noter que certains miARN, appelés mirtrons sont maturés sans être clivés par le complexe Drosha/DGCR8, mais sont directement excisés par le splicéosome (Okamura *et al.*, 2007). En effet, les extrémités de la structure tige boucle du pré-miARN correspondent précisément aux bornes d'épissage de l'intron.

Le produit de cette maturation (le pré-miARN) est un ARN double-brin en structure tige-boucle d'environ 70 nucléotides possédant une extrémité 3' saillante de 2 nucléotides, caractéristique de la digestion catalysée par les ribonucléases de type III. De plus, les activités exonucléases 5' \rightarrow 3' et 3' \rightarrow 5' induites par les clivages de l'endoribonucléase III Drosha dans un intron permet la dégradation de ce dernier avant même son excision, ce qui suggère que la dégradation nucléolytique induite par Drosha sur le pré-ARNm influence l'épissage ainsi que la maturation de nombreux ARNm (Morlando *et al.*, 2008).

2.2.2.2. Exportation cytoplasmique du pré-microARN

Suite à sa première maturation par Drosha, le pré-miARN qui est encore présent dans le noyau, est pris en charge par une protéine d'export nucléaire l'Exportine-5 (Figure 18) (Bohnsack *et al.*, 2004; Lund *et al.*, 2004). Cette dernière permet l'export du pré-miARN dans le cytoplasme de manière RanGTP-dépendante via les nucléopores (Bohnsack *et al.*, 2004; Lund *et al.*, 2004), mais aussi sa protection contre des digestions induites par des nucléases (Zeng & Cullen, 2004). La longueur de la séquence double brin et la présence de l'extrémité 3' saillante de 2 nucléotides sont des facteurs importants pour la reconnaissance et la prise en charge du pré-miARN par l'Exportine-5, contrairement à la séquence et à la structure de la boucle qui ne semblent pas impliquées dans le mécanisme d'export (Lund *et al.*, 2004; Zeng & Cullen, 2004).

2.2.2.3. Maturation cytoplasmique du pré-microARN

Une fois dans le cytoplasme, le pré-miARN est pris en charge par la seconde endoribonucléase III Dicer, au sein du complexe RLC (RISC-Loading Complex). Ce complexe contient une des quatre protéines AGO1-4 ainsi que les cofacteurs de Dicer, TRBP (Transactivation region (Tar) RNA Binding Protein) et PACT (Protein Activator of PKR) qui ne sont pas impliqués dans le clivage du pré-miARN mais qui facilitent l'assemblage du complexe RLC (Haase *et al.*, 2005; Lee *et al.*, 2006; Paroo *et al.*, 2009). En effet, la phosphorylation de TRBP par la MAPK/Erk (Mitogen-Activated Protein Kinase/Extracellular Regulated Kinase) observée durant la survie et la croissance cellulaires semble stabiliser le complexe et ainsi faciliter la maturation des miARN, montrant de ce fait l'importance de cette nouvelle voie de régulation dans la biogenèse des miARN (Paroo *et al.*, 2009).

La seconde maturation consiste au clivage par Dicer de l'extrémité possédant la structure boucle du pré-miARN, ce qui conduit à la production d'un ARN double brin d'environ 22 nucléotides correspondant au duplex miR-5P/miR-3P (Figure 18).

Par ailleurs, certains miARN, comme par exemple le miR-451 (Cheloufi *et al.*, 2010; Cifuentes *et al.*, 2010), peuvent être maturés sans être clivés par Dicer. En effet, la protéine AGO2, de par son activité endonucléasique, peut également entraîner le clivage du brin passager, ce qui semble faciliter la formation du complexe miRISC et la dissociation du duplex (Diederichs & Haber, 2007).

2.2.2.4. Sélection du brin mature du microARN

Durant ce processus, un des deux brins (brin guide) du duplex miR-5P/miR-3P est préférentiellement incorporé au sein du complexe miRISC, tandis que le second brin (brin passager ou miR*) est relâché puis dégradé (Sarnow *et al.*, 2006). Ainsi, la plupart des pré-miARN produisent un miARN mature dominant issu du brin guide, mais le brin passager peut également être incorporé dans le miRISC et agir tel un miARN fonctionnel.

La localisation du clivage par Drosha et Dicer détermine les bornes nucléotidiques 5' et 3' terminales des miARN. Aussi, il faut noter que de nombreuses isoformes de miARN avec des extrémités variables ont été identifiées du fait d'une maturation non uniforme. Ces variations peuvent avoir des conséquences importantes, notamment pour la borne 5', si elles induisent un décalage de la région « seed » essentielle à la fonctionnalité du miARN, entraînant une modification potentielle d'une part, des cibles, et d'autre part, de la sélection du brin guide. Généralement, le brin possédant l'extrémité 5' la moins stable au sein du duplex miR-5P/miR-3P est retenu. Cependant, il a également été décrit que l'interaction du nucléotide 5' terminal du miARN avec la protéine AGO2 peut aussi affecter directement l'efficacité de l'incorporation du miARN indépendamment de la stabilité des deux brins (Frank *et al.*, 2010). Ainsi, les propriétés intrinsèques du pré-miARN ou du duplex de miARN, mais aussi l'activité de différentes protéines accessoires peuvent induire des variations dans le ratio de miR-5P / miR-3P en fonction des tissus ou du stade de développement (Krol *et al.*, 2010).

2.2.2.5. Formation et régulation du complexe miRISC

Le complexe miRISC est un complexe ribonucléoprotéique (miRNP) contenant le miARN mature simple brin, une des quatre protéines Argonautes (AGO1-4) et le facteur GW182. Cependant, la présence ou non des autres facteurs du RLC, notamment Dicer et TRBP n'est pas clairement définie (Kawamata & Tomari, 2010). En effet, contrairement à Dcr-2 et R2D2 (Kawamata & Tomari, 2010) homologues de Dicer et TRBP chez la drosophile, la protéine Dicer humaine ne semble pas nécessaire à la prise en charge d'un siARN ni à l'action de celui-ci. De plus, une fois liée à un miARN mature, la protéine AGO semble se dissocier du reste du complexe (MacRae *et al.*, 2008).

Chez les mammifères, la répression de la synthèse protéique induite par les miARN ou les siARN via le complexe RISC dépend de la présence des protéines AGO1-4 et de la protéine GW182. Les protéines AGO sont caractérisées par la présence de trois domaines. Le domaine PAZ (Piwi-Argonaute-Zwille) qui forme un module spécifique de liaison à l'extrémité 3' saillante du miARN, le domaine MID (middle) contenant une poche d'aa fortement basiques qui se lie au phosphate situé en 5' des petits ARN et permet ainsi l'ancrage de ce dernier, et le domaine PIWI qui forme une structure similaire à la RNase H (Peters & Meister, 2007). Cependant, seule AGO2 possède un domaine PIWI réellement actif pouvant assurer le clivage de l'ARNm (Peters & Meister, 2007). Chez les vertébrés, trois gènes de la famille de GW182 ont été identifiés, GW182 ou TNRC6A (Trinucleotide repeat-containing protéines 6A), TNRC6B et TNRC6C. Les protéines correspondantes ont été récemment identifiées comme des facteurs indispensables au phénomène d'ARN interférence. Les protéines « GW182 » interagissent avec les protéines Argonaute uniquement après toutes les étapes de maturation et de prise en charge du miARN (Eulalio *et al.*, 2009). La fonction précise de ces protéines n'est pas très documentée, mais des études récentes ont montré qu'elles interagissent avec les PAIP (Poly(A) Binding Protein (PABP)-Interacting Proteins) permettant l'inhibition traductionnelle (Tritschler *et al.*, 2010).

La régulation des protéines AGO apparaît très importante au sein de la cellule. En effet, la surexpression de chacune des protéines AGO entraîne une augmentation du nombre de miARN matures au sein de cellules HEK293, suggérant que le taux d'AGO intracellulaire est un facteur limitant pour la formation du complexe RISC (Krol *et al.*, 2010). De plus, indépendamment de leur habilité intrinsèque à être intégrées au sein du miRISC, la variation de la concentration des protéines AGO entraîne une modification quantitative mais également qualitative des miARN associés au miRISC (Krol *et al.*, 2010). Différents mécanismes impliqués dans la régulation de la protéine AGO2 ont été identifiés comme par exemple, la protéine TRIM71 (tripartite motif-containing 71), qui induit une polyubiquitinylation d'AGO2 aboutissant à sa dégradation au sein du protéasome (Rybak *et al.*, 2009), les protéines HSP90 (Heat Shock Protein 90) et C-P4H(I) (type I collagen prolyl-4-hydroxylase) qui stabilisent AGO2 (Johnston *et al.*, 2010; Qi *et al.*, 2008) avec une action supplémentaire pour C-P4H(I) qui semble augmenter la localisation d'AGO2 au niveau des P-bodies (Processing Bodies), lieu de dégradation des ARNm.

2.3. Interaction microARN/ARNm

Deux modes d'action principaux sont imputés aux miARN : le clivage de l'ARNm cible ou le blocage de la traduction de cet ARNm cible. Cette dualité de mode d'action semble être fortement liée à la qualité de l'appariement du duplex miARN/ARNm. Une complémentarité parfaite entre le miARN et l'ARNm conduirait au clivage de ce dernier entre les nucléotides 10 et 11 du miARN, de manière comparable à un mécanisme de siARN, tandis qu'une complémentarité imparfaite conduirait au blocage de la traduction de cet ARNm et à une accélération de sa dégradation par déadénylation exonucléolytique.

Chez les métazoaires, bien que de rares cas de clivage de l'ARNm aient été observés comme par exemple pour le miR-196 chez les mammifères (Yekta *et al.*, 2004), la majorité des sites d'interaction étudiés correspondent à des complémentarités imparfaites entre les deux ARN, et plus précisément, entre la région « seed » du miARN et la 3'UTR de l'ARNm. A l'inverse, chez les plantes, bien que le processus d'inhibition de la traduction soit également présent, la plupart des interactions correspondent à des complémentarités souvent parfaites aboutissant au clivage de l'ARNm (Brodersen *et al.*, 2008; Voinnet, 2009).

La régulation par les miARN dépend de nombreux facteurs, dont principalement la complémentarité du miARN pour le site cible mais également l'environnement de la cible, les interactions avec des protéines de liaison à l'ARN, ou encore la position de la cible au sein de l'ARNm.

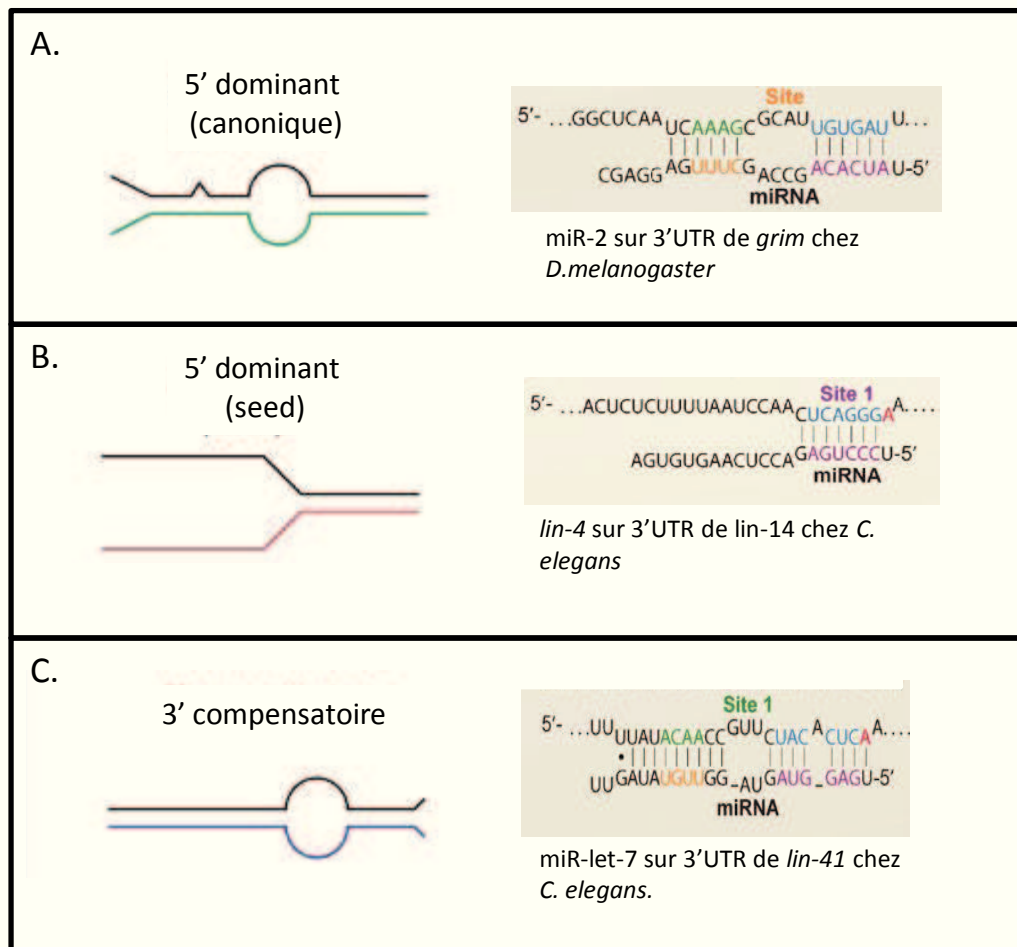


Figure 19 : Représentation des trois types d'interaction fonctionnelle entre un microARN et sa cible (d'après Brennecke et al., 2005).

(A) Exemple du ciblage 5'dominant canonique avec une complémentarité de l'intégralité de la séquence « seed » et des complémentarités en 3' du miARN. (B) Exemple du ciblage 5'dominant présentant une complémentarité de la totalité de la séquence seed, mais pas de complémentarités en 3' du miARN. (C) Exemple d'un ciblage 3'compensatoire, qui ne présente pas de complémentarité parfaite de la séquence « seed » mais un fort appariement en 3' permettant de compenser le mésappariement de la « seed ».

2.3.1. Reconnaissance de l'ARNm par le microARN

L'étude des appariements entre le miARN et sa cible par l'analyse des interactions endogènes ainsi que par des expériences de mutagenèse ont permis de mettre en lumière l'importance de chaque nucléotide du miARN dans le processus de reconnaissance de l'ARNm cible. Le premier nucléotide du miARN, qui est souvent une uridine, est impliqué dans la liaison aux protéines Argonaute et ne semble pas interagir directement avec l'ARNm cible durant le processus de reconnaissance (Ma *et al.*, 2005; Parker *et al.*, 2005). De plus, il a été montré que la séquence « seed » correspondant aux nucléotides 2 à 8 du miARN était essentielle dans ce processus de reconnaissance. En effet, des études suggèrent l'existence d'une étape préliminaire dite de nucléation, favorisant la liaison ultérieure du miARN sur sa cible (Ma *et al.*, 2005; Parker *et al.*, 2005; Wang *et al.*, 2008b). Ce modèle propose une structuration spatiale des 7 nucléotides de la « seed » par le complexe RISC afin de ressembler à une demi-hélice de type A qui permettrait d'augmenter l'affinité et la spécificité pour l'ARNm ciblé (Bartel, 2009; Wang *et al.*, 2008b). Par ailleurs, Brennecke et ses collaborateurs (Brennecke *et al.*, 2005) avaient défini deux principales catégories d'interactions miARN/ARNm (Figure 19). La première, correspondant à un appariement principal en 5' du miARN dit « 5' dominant », où la région « seed » est l'élément principal de reconnaissance (Figure 19A et B). La seconde catégorie dite « 3' compensatoire » regroupe les interactions ne présentant pas de complémentarité parfaite dans la région « seed » mais possédant un fort appariement en 3' du miARN permettant de compenser le mésappariement en 5' (Figure 19C). De plus, entre ces 2 catégories d'interaction extrêmes de nombreuses interactions miARN/ARNm présentent des appariements supplémentaires en 3' du miARN en plus d'une complémentarité parfaite de la « seed ». Bien que ces trois types d'interactions soient fonctionnels, il semble que les interactions par sites 5' dominant soient observées majoritairement.

Récemment, le modèle a été précisé afin d'intégrer les observations précédentes (Bartel, 2009). Ainsi, lors d'une complémentarité imparfaite entre le miARN et sa cible des appariements additionnels de la région du miARN située en 3' de la « seed » peuvent s'effectuer après la nucléation, notamment au niveau des nucléotides 13 à 16 du miARN, permettant l'incorporation de l'ARNm en un petit segment en hélice sans perturbation majeure de la structure de la protéine AGO ou du miARN (Figure 20B et 20E). Il est intéressant de noter que dans ce mode de reconnaissance, le miARN et l'ARNm ne s'enroulent pas totalement entre eux contrairement à la situation présumée lors d'une complémentarité parfaite (Figure 20C et 20D). Lors de complémentarités parfaites, les auteurs proposent que la propagation de la nucléation au sein de la région centrale conduise à un réarrangement conformationnel de la région 3' du miARN qui entraînerait la perte partielle de la liaison à AGO, rendant possible l'enroulement du miARN et de l'ARNm (Figure 20). Par la suite, la

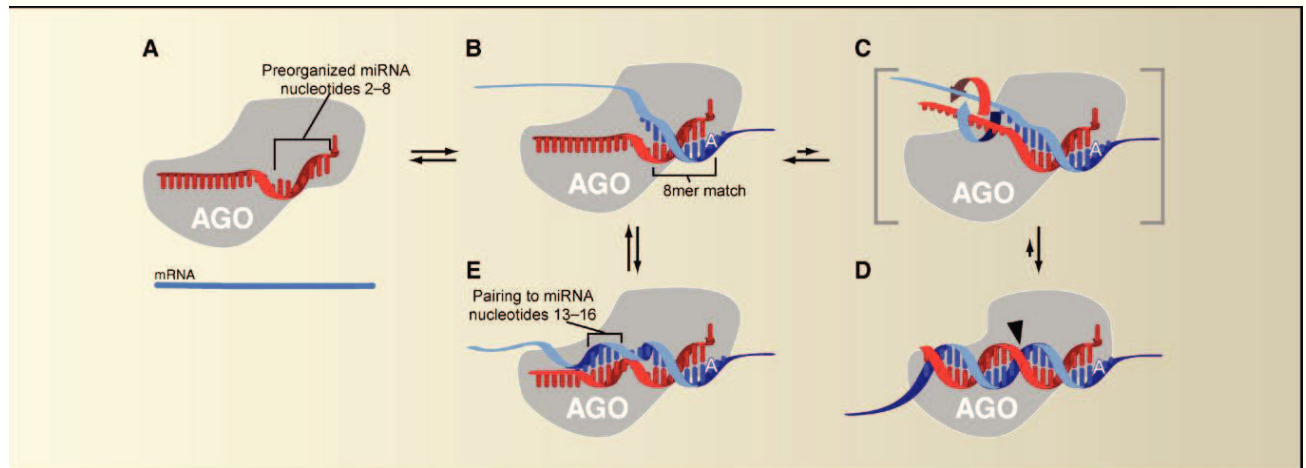


Figure 20 : Modèle des interactions des différentes régions d'un miARN avec sa cible (Bartel et al., 2009).

(A) Le miARN, représenté en rouge, est lié à la protéine Argonaute et sa séquence seed (nucléotides 2-8) est organisée structurellement pour favoriser l'interaction avec l'ARNm, représenté en bleu. (B) Reconnaissance de 8 nucléotides de l'ARNm par la poche de liaison préformée par la région seed. Le premier nucléotide de la cible semble être pris en charge par AGO ou une autre protéine du complexe RISC. (C) Réarrangement conformationnel supposé de la partie 3' du miARN, retrouvé lors des rares cas de complémentarité parfaite du miARN endogène avec l'ARNm ou dans le cas des siARN. Suite à la propagation de la nucléation au sein de la région centrale, la protéine AGO perd sa liaison à la région 3' du miARN, permettant ainsi l'enroulement du miARN et de l'ARNm. (D) La protéine AGO se referme sur le duplex ce qui place le site actif en position pour le clivage de l'ARNm. (E) Dans les cas d'appariement additionnels en 3' du miARN, l'ARNm peut se lier aux nucléotides 13 à 16 du miARN, s'incorporant en un petit segment en hélice mais sans perturbation majeure de la protéine AGO ou du miARN.

protéine AGO se refermerait sur le duplex, afin de positionner le site actif pour le clivage de l'ARNm. Par ailleurs, il est intéressant de noter que la séquence « seed » ainsi que les cibles complémentaires de cette séquence « seed » sont très conservées entre les différents organismes.

2.3.2. Impact de la position du site cible

Il a longtemps été admis que les miARN n'agissaient que sur la région 3'UTR des ARNm, et ce n'est que plus récemment que des cibles situées dans la séquence codante ou dans la région 5' UTR de l'ARNm ont été validées. De ce fait, chez les métazoaires, la majorité des sites étudiés sont localisés dans la 3'UTR des ARNm, et l'importance de la fixation des miARN dans les 5' UTR et les séquences codantes reste peu caractérisée. Une étude récente a montré par un système basé sur la reconnaissance de séquences conservées relatives aux séquences « seed » des miARN, la présence d'un grand nombre de sites « cible » potentiels au niveau des ORF dans les génomes de la drosophile et de l'homme (Schnall-Levin *et al.*, 2010). Chez la drosophile, le nombre de sites conservés semble égal entre les régions 3'UTR et les ORF mais est significativement plus faible au sein des régions 5'UTR. Par contre, chez l'homme, ces sites semblent plus présents au sein des régions 3'UTR qu'au sein des ORF et plus rares dans les 5'UTR (Schnall-Levin *et al.*, 2010).

Au sein même de la région 3'UTR, des préférences de localisation ont été observées. En effet, il semblerait qu'au sein de longues régions 3'UTR, des sites situés aux extrémités de l'ARNm, seraient plus propices à une inhibition par le miARN (Grimson *et al.*, 2007). De plus, un effet synergique entre plusieurs sites relativement proches au sein des régions 3'UTR a aussi été mis en évidence (Grimson *et al.*, 2007).

2.3.3. Influence de la structure de l'ARNm

Le contexte nucléotidique dans lequel se situe le site de liaison est aussi crucial sur l'effet de la répression (Kertesz *et al.*, 2007). En effet, en prenant en compte l'accessibilité du miARN pour sa cible mais également l'accessibilité du complexe miRISC au niveau de ces sites, des miARN peuvent perdre leur affinité pour une cible, si cette dernière présente un contexte défavorable, résultant en une perte de leur pouvoir inhibiteur. Bien que la corrélation ne soit pas établie, il est intéressant de noter que les 30 nucléotides situés en amont et en aval des cibles fonctionnelles sont significativement plus riches en AU que les sites non fonctionnels (Grimson *et al.*, 2007), laissant supposer que des structures plus stables induites par de plus forts taux en guanine et cytosine empêcheraient la reconnaissance de la cible.

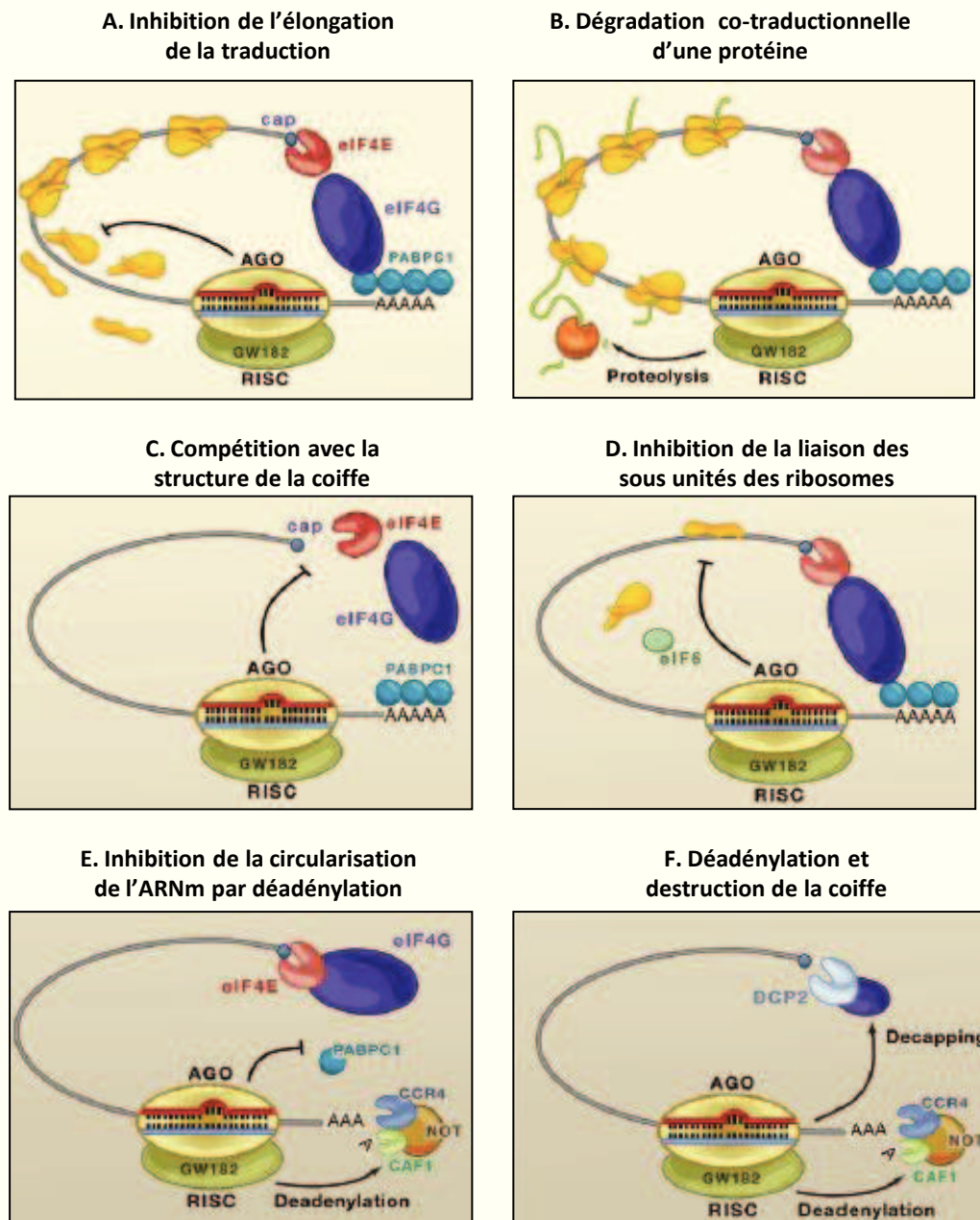


Figure 21 : Représentation des différents mécanismes d'inhibition de la traduction par les microARN (Eulalio et al., 2008).

Le complexe miRISC est représenté sous la forme d'un complexe minimal incluant une protéine Argonaute (AGO ; en jaune) et le facteur GW182 (en vert). L'ARNm est représenté sous une configuration de boucle fermée suite aux interactions entre la PABPC1 (cytoplasmic polyA binding protein) située au niveau de la queue polyA et le facteur eIF4E (eucaryotic translation initiation factor 4E) situé au niveau de la coiffe. (A) Le miARN, représenté en rouge, réprime la traduction de l'ARNm cible par blocage de l'élongation de la traduction ou par dissociation précoce des ribosomes (ribosome drop-off). (B) Dégradation co-traductionnelle de la chaîne polypeptidique naissante. (C) Compétition pour la liaison avec la structure de la coiffe entre la protéine Argonaute et le facteur eIF4E. (D) Recrutement du facteur eIF6 prévenant la liaison entre la grande sous unité et la petite sous unité du ribosome. (E) La protéine Argonaute empêche la formation de la boucle de l'ARNm par un mécanisme mal caractérisé mais incluant la déadénylation. (F) Le miARN déclenche la déadénylation et une dégradation significative de la coiffe de l'ARNm ciblé.

2.3.4. Influence des protéines de liaison à l'ARN

La présence de protéines de liaison à l'ARN à proximité de la cible d'un miARN peut induire soit une coopération avec le complexe miRISC pour l'inhibition de la traduction, soit une compétition pour l'accessibilité au site de liaison. Chez *C.elegans*, la protéine Puf-9 (PUF (Pumilio/FBF) domain-containing 9) se fixerait à proximité de cibles du miARN let-7 au sein de la région 3' UTR de *Lin-41* et de *hbl-1* et elle serait indispensable à l'inhibition induite par let-7 (Nolde *et al.*, 2007). Par ailleurs, il a été montré que dans certaines conditions de stress, l'inhibition de la traduction de l'ARNm du gène *CAT-1* (Cationic amino Acid Transporter 1) par miR-122 pouvait être supprimée par la liaison de la protéine HuR à l'ARNm correspondant (Bhattacharyya *et al.*, 2006). De même, la protéine Dnd-1 (Dead end 1), inhibe l'action du miR-372 et du miR-221 sur leur cible respective *LATS2* et *p27* dans des cellules tumorales humaines, ainsi que l'action du miR-430 sur les régions 3'UTR de *nanos1* et *TDRD7* au sein de cellules germinales de poisson zèbre (Kedde *et al.*, 2007). Ces 2 protéines inhibitrices de l'action des miARN semblent se fixer sur des séquences spécifiques riches en U pour Dnd-1 (Kedde *et al.*, 2007) ou en résidus AU pour HuR (Bhattacharyya *et al.*, 2006), situées à proximité des cibles. Ainsi, bien que le nombre d'exemples d'implication des protéines de liaison à l'ARN dans la régulation de l'ARN interférence reste encore limité, ceci suggère qu'il pourrait correspondre à un mécanisme de régulation de l'inhibition relativement généralisé au sein des miARN.

2.4. Mécanismes régulés par les microARN

Du fait que la majorité des cibles identifiées à ce jour soient présentes au sein des régions 3'UTR des ARNm, l'action précise des miARN dans leur complexe miRISC a été principalement caractérisée au sein de cette région, tandis que leur action lors de sa fixation dans les régions codantes ou dans les régions 5'UTR est encore mal connue. Au sein de la région 3'UTR, les miARN peuvent agir par différents mécanismes distincts lors de l'initiation de la traduction ou au cours de la traduction (Figure 21) (Eulalio *et al.*, 2008; Fabian *et al.*, 2010). Cependant, ces mécanismes d'actions restent encore controversés. De plus, chez les métazoaires, les miARN semblent induire une dégradation significative des ARNm malgré une complémentarité imparfaite entre les deux ARN, par séquestration des ARNm au sein des P-bodies, où ils subissent une destruction de leur queue polyA et de leur coiffe.

2.4.1. Mécanismes d'inhibition de l'initiation de la traduction

Différents modes d'action des miARN conduisant au blocage de l'initiation de la traduction ont été recensés. Tout d'abord, le complexe peut entrer en compétition avec la coiffe de l'ARNm puisqu'une diminution ou un arrêt de l'inhibition induite par certains miARN sur des ARNm possédant une coiffe non fonctionnelle (ApppG) ou pilotés par une séquence IRES ont été mis en évidence (Humphreys *et al.*, 2005; Pillai *et al.*, 2005). Dans ce cas, le complexe miRISC pourrait inhiber la formation de la structure en boucle entre la coiffe et la queue polyA de l'ARNm réalisée par le complexe eIF4F. Il semblerait que la protéine AGO puisse interagir avec le facteur eIF4E du complexe eIF4F en empêchant sa liaison avec la coiffe (figure 21C) ou en interagissant avec la queue polyA. Dans ce dernier cas, le complexe miRISC pourrait interagir directement avec la PABP (polyA binding protein) empêchant sa liaison à eIF4G, ou alors recruter via GW182 des 3'-5' exoribonucléases, telles que le complexe CCR4/NOT1 (Carbon Catabolite Repression 4 / Negative On TATA less 1) qui comprend les déadénylases CCR4/CNOT6 et CAF1/CNOT7, responsables de la déadénylation de l'ARNm (Figure 21E).

Par ailleurs, lors d'essais *in vitro* (Wang *et al.*, 2008a), il a été montré que les miARN pouvaient également bloquer l'initiation de la traduction en empêchant la liaison entre la petite sous-unité ribosomale 40S et la grande sous-unité 60S (Figure 21D). Suite à la co-immunoprécipitation du facteur eIF6 avec AGO2/Dicer/TRBP, il a été proposé que le complexe miRISC puisse recruter ce facteur inhibiteur de la jonction des sous-unités des ribosomes (Chendrimada *et al.*, 2007).

2.4.2. Mécanismes d'inhibition post-initiation de la traduction

Différentes études ont mis en évidence une inhibition de la traduction après son initiation par identification de miARN et de certains composants du complexe miRISC au sein de polysomes (Maroney *et al.*, 2006; Nottrott *et al.*, 2006). De même, par observation de répression de la traduction sur des protéines pilotées par des IRES, Petersen et ses collaborateurs (Petersen *et al.*, 2006) ont montré grâce à un gène rapporteur synthétique comprenant six cibles de miARN partiellement complémentaires d'un siARN que ce dernier induisait une dissociation prématurée des ribosomes (ribosome drop-off) aboutissant à des polypeptides se terminant prématurément (Figure 21A). Par ailleurs, suite à la traduction effective d'ARNm ciblés par des miARN sans que la protéine correspondante ne puisse être détectée (Nottrott *et al.*, 2006), les auteurs ont proposé un mécanisme de dégradation co-traductionnelle de la chaîne polypeptidique naissante par une protéase encore inconnue (Figure 21B), indépendamment du protéasome (Nottrott *et al.*, 2006).

2.4.3. Dégradation de l'ARNm

En plus de l'effet sur l'inhibition de la traduction, les miARN induisent aussi une séquestration au sein de compartiments cytoplasmiques appelés P-bodies (ou GW-bodies), correspondant au lieu de dégradation des ARNm (Eulalio *et al.*, 2007). En effet, de nombreux miARN, des protéines AGO et des ARNm ciblés par ces miARN ont été identifiés dans ces compartiments cytoplasmiques qui contiennent de fortes concentrations de protéines GW182, de déadénylases et de protéines permettant la dégradation de la coiffe telles que DCP1 et DCP2 (Figure 21F) (Behm-Ansmant *et al.*, 2006; Rehwinkel *et al.*, 2005).

2.4.4. Activation de la traduction

Le rôle des miARN ne semble pas se limiter à l'inhibition de l'expression génique et quelques cas favorisant l'expression d'un gène ont été identifiés. En effet, les miARN semblent pouvoir activer la transcription d'un gène en se fixant dans sa région promotrice (Place *et al.*, 2008). Par ailleurs, plusieurs cas d'augmentation de la traduction d'un gène ont également été répertoriés notamment avec la fixation du miR-10a dans la région 5'UTR de protéines ribosomales (Orom *et al.*, 2008), et celle du miR-122 près de l'IRES du virus de l'hépatite C (Henke *et al.*, 2008). De même, les miR let-7 semblent activer la traduction de certains ARNm dans des cellules quiescentes en absence de sérum (Vasudevan *et al.*, 2007; 2008).

2.5. MicroARN et virus

2.5.1. Les virus encodant des microARN

Les miARN viraux ont été identifiés pour la première fois en 2004 chez le gammaherpesvirus HHV-4 (Pfeffer *et al.*, 2004). A l'heure actuelle, des miARN ont été mis en évidence chez des virus appartenant à cinq familles différentes : les Polyomavirus (Sullivan *et al.*, 2005), les Adenovirus (Aparicio *et al.*, 2006), les Ascovirus (Hussain *et al.*, 2008), les Baculovirus (Singh *et al.*, 2010) et les Herpesvirus (Pfeffer *et al.*, 2004). C'est au sein de cette dernière famille qu'a été identifiée la majorité des miARN viraux, et ce, dans les trois sous-familles α , β et γ *herpesvirinae*.

La voie de biogénèse des miARN viraux est assurée par la cellule hôte et le mécanisme d'action ainsi que le mode de répression semblent être régis par les mêmes règles que pour les miARN cellulaires. Ils sont présents dans des introns et des exons de gènes codant des protéines, ou dans le cas du murine herpesvirus de type 68 (MHV-68), au sein d'ARNt (Pfeffer *et al.*, 2005; Reese *et al.*, 2010).

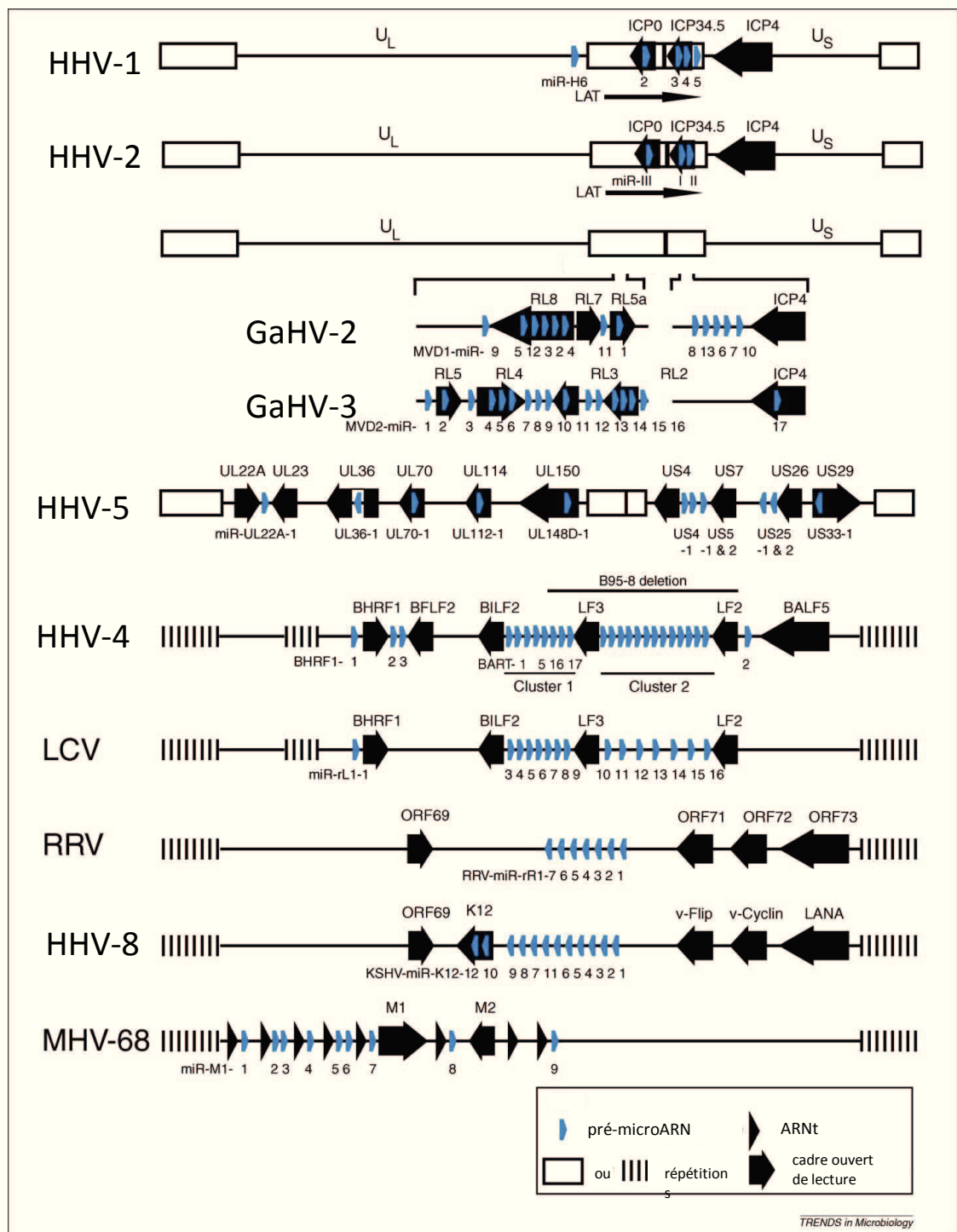


Figure 22 : Localisation des microARN au sein des génomes de différents herpesvirus (Boss et al., 2009). Les génomes des quatre alphaherpesvirus HHV-1 (Human Herpesvirus-1), HHV-2 (Human Herpesvirus-2), GaHV-2 (Gallid Herpesvirus 2), GaHV-3 (Gallid Herpesvirus 3), du betaherpesvirus HHV-5 (Human Herpesvirus-5) et des cinq gammaherpesvirus HHV-4 (Human Herpesvirus-4), LCV (Lymphocryptovirus), RRV (Rhesus Rhadinovirus), HHV-8 (Human Herpesvirus-8), MHV-68 (Murine gammaherpesvirus type 68) sont représentés. Les tailles ne sont pas représentatives des tailles réelles des génomes. Abréviations : (U_S) région unique courte, (U_L) région unique longue, (LAT) Latency Associated Transcript.

Il faut remarquer que les miARN sont relativement spécifiques des espèces virales, au sein desquelles les séquences sont très conservées d'une souche à l'autre. De plus, ils sont généralement codés sous la forme de clusters de miARN (Figure 22) dont l'expression est souvent spécifique du stade viral.

Du fait de la maturation nucléaire et cytoplasmique des miARN, la présence des miARN viraux semble spécifique des virus à ADN à réplication nucléaire. Dans le cas de HIV-1, des miARN ont été identifiés, cependant leur existence reste très controversée (Klase *et al.*, 2009; Lin & Cullen, 2007; Omoto *et al.*, 2004; Pfeffer *et al.*, 2005). Il faut noter qu'une étude récente a montré qu'un miARN présent artificiellement chez Tick-Borne Encephalitis Virus, un virus à ARN à réplication cytoplasmique, peut être maturé et aboutir à un miARN fonctionnel (Rouha *et al.*, 2010).

Bien que les miARN cellulaires ne semblent pas spécifiquement exprimés pour contrer un virus particulier, certains agissent négativement sur la réplication virale et sont impliqués dans la défense antivirale. Afin de contrecarrer leur action, différentes protéines ou ARN viraux ont été décrits pouvant interférer avec la machinerie de biosynthèse des miARN cellulaires (Bivalkar-Mehla *et al.*, 2010). Par ailleurs, il est intéressant de noter que la plupart des infections virales entraîne des variations dans les profils d'expression de certains miARN cellulaires suite à des dérégulations positives ou négatives induites par la cellule ou par le virus.

A ce jour, peu de miARN viraux avec cibles identifiées ont été validées pour les miARN viraux mais les premiers résultats semblent indiquer que les miARN viraux agiraient principalement en ciblant leurs propres gènes afin de réguler leur cycle viral ou en ciblant des gènes cellulaires pour contourner le rôle de ces derniers dans la réponse immunitaire, le cycle cellulaire et/ou l'apoptose.

2.5.2. Régulation du cycle viral par les microARN viraux

La régulation du cycle viral par des miARN viraux semble être un mécanisme courant chez les virus encodant des miARN.

Chez les polyomavirus, le miARN possède une fonction conservée entre les différentes espèces virales (Seo *et al.*, 2008). En effet, ce miARN présent sur le brin anti-sens de la région 3'UTR de l'antigène T, est exprimé lors de la phase tardive de l'infection. Il permettrait la régulation du taux d'antigène T qui est une cible des cellules T cytotoxiques, afin de contrôler l'échappement au système immunitaire et surtout il permettrait l'expression des gènes viraux tardifs (Sullivan *et al.*, 2005).

Comme chez les herpesvirus les miARN ne sont pas conservés, les processus observés sont par essence spécifiques des trois sous-familles virales. Pour l'alphaherpesvirus HHV-1, quatre des six miARN (hsv1-miR-H2, H3, H4 et H5) sont codés au sein du transcrit LAT, un long ARNnc exprimé spécifiquement durant la phase de latence (Figure 23). Parmi ces miARN, hsv1-miR-H2 correspond au brin anti-sens du gène *ICP0* (transactivateur essentiel du cycle viral) tandis que les miARN hsv1-miR-H3 et hsv1-miR-H4 au brin anti-sens de *ICP34.5* (gène clef de la neurovirulence), permettant le

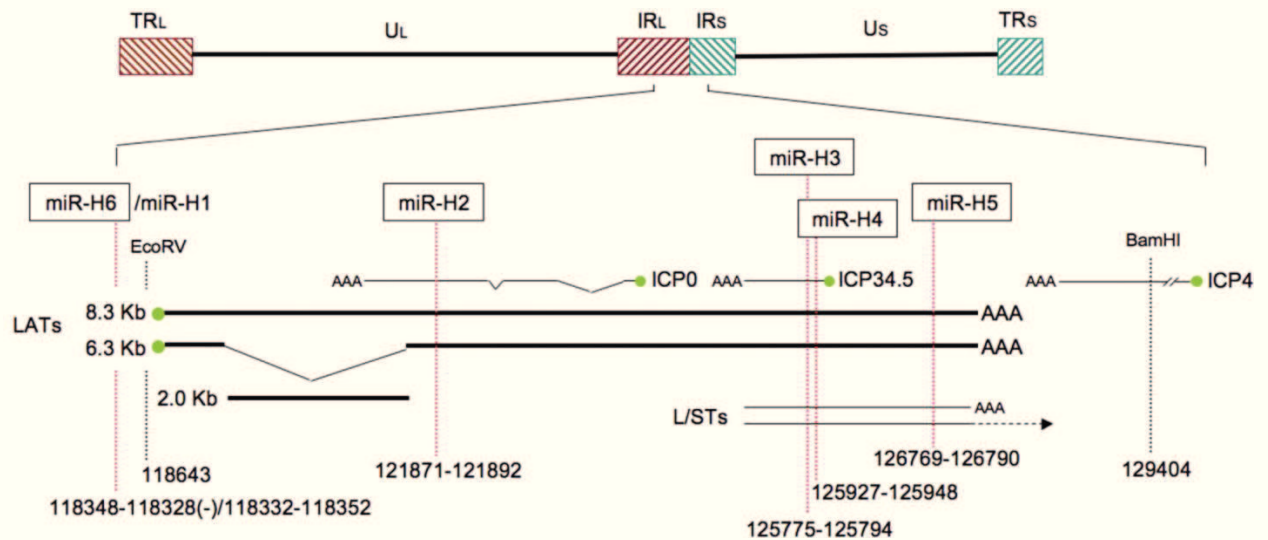


Figure 23 : Localisation des miARN de HHV-1 au sein du locus LAT (Umbach *et al.*, 2008).

La localisation, l'orientation ainsi que la taille relative des différents transcrits de la région sont indiquées. Les coordonnées des différents miARN et des sites de restriction sont indiqués en fonction de la souche 17 syn+ (NC_001806) de HHV-1. A l'exception de miR-H6, tous les miARN sont dans la même orientation que le gène LAT. Abréviations : TR : Répétition Terminale ; IR : Répétition Interne ; U_L : Région Unique Longue ; U_S : Région Unique Courte.

ciblage avec une complémentarité parfaite de ces deux gènes impliqués dans le cycle lytique. Chez HHV1, parmi ces trois miARN, hsv1-miR-H2 est le seul à avoir été investigué et dont le rôle dans la diminution efficace de la synthèse d'ICP0 a été validé (Umbach *et al.*, 2008). De plus, les miARN homologues chez HHV-2 présentant une localisation identique mais des séquences différentes, inhibent la synthèse d'ICP34.5 et d'ICP0 (Tang *et al.*, 2009). Par ailleurs, le miARN hsv-1-miR-H6, qui est exprimé par un autre transcrit à proximité du LAT (Figure 23) est capable également d'inhiber la synthèse du transactivateur ICP4 par fixation dans sa région 3'UTR via une complémentarité imparfaite (Umbach *et al.*, 2008). Ainsi, les miARN codés dans la région du LAT de HHV-1 et 2, semblent jouer un rôle primordial dans le maintien de la latence, par inhibition des transcrits essentiels à la réplication lytique. Chez les betaherpesvirus, et plus spécifiquement chez HHV-5 (HCMV ; Human Cytomegalovirus) hcmv-miR-UL112 est le miARN le mieux caractérisé. Ce miARN est capable d'inhiber l'expression de différents gènes de la région MIE (Major Immediate Early), dont la protéine IE72 et potentiellement les gènes U_L120/121 et U_L112/113, impliqués dans la réplication virale (Grey *et al.*, 2007; Murphy *et al.*, 2008). Grey et ses collaborateurs ont observé un effet de la présence des séquences cibles de ces miARN dans les régions 3' UTR de ces trois gènes sur la réplication virale (Grey *et al.*, 2007), cependant, via l'utilisation de virus recombinants n'exprimant pas le miARN ou dont la cible de IE72 était muté, Murphy et ses collaborateurs ont confirmé un effet du miARN sur IE72 mais n'ont pas observé de différences entre la réplication des virus natifs et mutés (Murphy *et al.*, 2008). Pour finir, parmi les gammaherpesvirus et les mécanismes d'inhibition de la réplication virale, il a été identifié chez HHV-4, ebv-miR-BART2 qui cible le gène viral BALF5 codant pour la polymérase (Barth *et al.*, 2008; Pfeiffer *et al.*, 2004) permettant aussi le maintien de la latence virale de HHV-4.

2.5.3. Régulation des réponses cellulaires par des microARN viraux

Le second impact principal des miARN viraux concerne l'inhibition de protéines impliquées dans la réponse immunitaire cellulaire innée ou adaptative, dans la survie cellulaire et la prolifération. Ces mécanismes d'échappement de la reconnaissance cellulaire ont été mis en évidence par modulation de protéines cellulaires mais également de protéines virales.

Ainsi, il été démontré que le miARN ebv-miR-BART5 d'HHV-4 cible le gène cellulaire PUMA codant pour une protéine pro-apoptotique, suggérant que ce miR viral participe à la mise en place de la latence en assurant la survie des cellules infectées (Choy *et al.*, 2008). De plus, trois autres miARN du cluster BART (ebv-miR-BART-16, 17-5P et 1-5P) sont capables de moduler l'expression de l'oncoprotéine virale LMP1 (Latent Membrane Protein 1) permettant la survie cellulaire par inhibition de l'apoptose induite par LMP1 durant la latence (Lo *et al.*, 2007). Toujours chez HHV-4, le cluster de miARN BHRF1 est impliqué dans l'inhibition de l'apoptose et favorise la progression du

cycle cellulaire et la prolifération durant la phase précoce de l'infection de cellules B primaires humaines (Seto *et al.*, 2010). De plus, ebv-miR-BHRF1-3, qui est principalement exprimé durant la phase de latence III, est capable d'inhiber l'expression de la chimiokine CXCL-11, impliquée dans l'attraction de cellules T, conduisant à un échappement à la surveillance cellulaire (Xia *et al.*, 2008).

Le gène THBS1 (Thrombospondin1), un gène suppresseur de tumeur mais aussi un facteur anti-angiogénique fortement dérégulé durant la latence virale de HHV8 possède 34 séquences cibles potentielles pour les miARN viraux au sein de sa région 3'UTR (Samols *et al.*, 2007). Enfin, chez HHV-5, Stern-Ginossar et ses collaborateurs ont montré qu'un des miARN ciblant les MIE : le miARN hcmv-miR-UL112, était capable de cibler la protéine cellulaire MICB (Major histocompatibility complex class I polypeptide-related sequence B). Cette protéine essentielle à la reconnaissance des cellules infectées par les cellules NK (Natural Killer) représente le premier rempart de la défense immunitaire innée (Stern-Ginossar *et al.*, 2007). Il est intéressant de noter que dans ce modèle, la protéine virale U_L16 inhibe également les fonctions de la protéine MICB permettant potentiellement un effet synergique afin de protéger les cellules infectées (Cosman *et al.*, 2001).

2.6. Rôle des ADAR dans l'ARN interférence

Durant ces dernières années, un lien entre les protéines ADAR et la régulation de l'ARN interférence a été observé pour plusieurs miARN. En effet, l'édition par les protéines ADAR pourrait réguler la biogenèse des miARN ainsi que leur spécificité.

2.6.1. Régulation de la biogenèse des microARN

A titre d'exemple, jusqu'à 11 désaminations A→I induites par ADAR1 et 2 ont été décrites pour le pri-miR-142 entraînant une altération de sa structure tige-boucle et empêchant ainsi sa maturation par Drosha (Yang *et al.*, 2006). De plus, ces auteurs ont montré *in vitro* que la transformation des appariements A:U en mésappariement IxU conduisait à la reconnaissance et au clivage de ce miARN par la protéine Tudor SN (Tudor Staphylococcal Nucléase) (Yang *et al.*, 2006). En effet, cette protéine retrouvée associée à de nombreuses protéines du complexe RISC, possède une activité ribonucléase spécifique des ARNdb contenant des inosines (Scadden & O'Connell, 2005), qui entraînerait la destruction du pri-miARN édité (Yang *et al.*, 2006).

Dans le même esprit, l'édition d'un pri-miARN peut également inhiber sa maturation par la protéine Dicer (Kawahara *et al.*, 2007a). En effet, l'édition de 2 sites spécifiques du pri-miR-151 par ADAR1 n'empêche pas sa maturation en pré-miR-151 édité par Drosha. Cependant, ce dernier ne sera

pas maturé en miR-151 par Dicer aboutissant à son accumulation sous la forme éditée dans le cytoplasme (Kawahara *et al.*, 2007a).

Par ailleurs, l'exemple des miARN du cluster miR-376 permet de souligner la complexité de la régulation des miARN par les protéines ADAR au sein d'un même cluster. Alors que les deux brins du miR-376a1 seraient préférentiellement édités par ADAR2, le pri-miR-376a2 est édité dans sa partie 5P par ADAR2 et dans sa partie 3P par ADAR1 (Heale *et al.*, 2009). De plus, ADAR2 semblerait également impliquée dans l'inhibition de la maturation des miR-376a1 et a2 par Drosha via un mécanisme indépendant de son activité d'édition (Heale *et al.*, 2009).

Pour terminer ce paragraphe avec un exemple viral, récemment, Iizasa et ses collaborateurs ont observé que le transcrit primaire de ebv-miR-BART6 (pri-miR-BART6) était édité dans des cellules transformées par HHV4 (Iizasa *et al.*, 2010). Cette désamination A→I du 20^{ème} nucléotide de ebv-miR-BART6-3P entraînerait le blocage du clivage du pri-miR-BART6 par Drosha (Iizasa *et al.*, 2010) empêchant les étapes de maturation ultérieures et notamment l'incorporation de ebv-miR-BART6 dans le complexe RISC. Ainsi, ebv-miR-BART6-5P qui est le brin majoritairement incorporé au sein du complexe RISC, ne pourrait plus cibler l'oncogène viral EBNA2 qui est nécessaire à la transition en latence de type III (Iizasa *et al.*, 2010).

2.6.2. Régulation de la spécificité des microARN

Un accroissement potentiel de la gamme des cibles reconnues par des miARN par édition de la séquence « seed » a été rapporté pour la plupart des 7 miARN du cluster miR-376. Plus particulièrement, les auteurs ont montré que l'édition d'un nucléotide de la « seed » du miR-376a1-5P par ADAR2 entraînait un changement de cibles (Kawahara *et al.*, 2007b). Enfin, bien que non encore mise en évidence, la régulation de l'interférence pourrait également se faire par édition de la cible au niveau de l'ARNm. En effet, la région 3'UTR des messagers portant de nombreuses cibles de miARN est assemblée en structure tige boucle permettant une édition par ADAR et un échappement au contrôle par certains miARN suite à une perte potentielle de la cible.

3. Le virus de la Maladie de Marek

3.1. La maladie de Marek

3.1.1. Historique

La Maladie de Marek (MD : Marek's Disease) a été décrite pour la première fois chez les poulets fermiers en 1907 par un vétérinaire hongrois nommé Jozsef Marek (Marek, 1907). Cette maladie hautement contagieuse à répartition mondiale se caractérise par une infiltration de cellules lymphocytaires au niveau des nerfs périphériques et par le développement de tumeurs lymphoïdes au niveau des organes viscéraux, des muscles et de la peau (Nazerian, 1973). Au cours des années 1950, alors que la sévérité de la maladie de Marek augmentait régulièrement avec l'industrialisation de la filière avicole aux Etats-Unis et en Europe, une forme aiguë de la maladie est apparue (Benton, 1957). Cette nouvelle forme, associée à un fort taux de mortalité (30-60%) chez les animaux infectés, est devenue la forme prédominante de la maladie au début des années 1960. Dès lors, les travaux de recherche portant sur la maladie de Marek s'étant intensifiés, l'agent étiologique fut isolé quelques années plus tard, par deux équipes indépendantes à partir de co-cultures de cellules sanguines ou tumorales de poulets et de cellules de reins de poulets (Churchill & Biggs, 1967) ou de fibroblastes embryonnaires de canard (Nazerian *et al.*, 1968). Le virus, reconnu comme étant un herpesvirus, est alors appelé Virus de la maladie de Marek (MDV : Marek's Disease Virus rebaptisé récemment GaHV-2 : Gallid Herpesvirus de type 2). L'impact épidémiologique et économique de la maladie n'a cessé d'augmenter jusqu'à la découverte d'un vaccin élaboré à partir de la souche virulente HPRS-16 atténuée par passages successifs en culture de cellules (Churchill *et al.*, 1969). Ainsi, la mise en place de la vaccination a permis de maîtriser et de stabiliser l'impact économique de la maladie de Marek, qui est le premier modèle de maladie pour lequel une néoplasie fut contrôlée de manière satisfaisante par la vaccination. Cependant, depuis la fin du 20^{ème} siècle des souches de plus en plus virulentes, dites « hyper virulentes » (vv et vv+) sont apparues, entraînant des échappements à la vaccination.

3.1.2. Transmission

Bien qu'infectant principalement les poulets (*Gallus gallus*), le virus de la maladie de Marek est aussi capable d'infecter d'autres espèces d'oiseaux telles que la caille, la dinde ou le faisan. GaHV-2 se transmet par voie aérienne, par contact direct ou indirect entre les poulets. Son mode de

transmission par production de formes virales libres au niveau des follicules plumeux a été montré pour la première fois par l'équipe de Calnek en 1970 (Calnek *et al.*, 1970). Ainsi, le virus se multiplie fortement au niveau des kératinocytes des follicules plumeux à partir de la deuxième semaine suivant l'infection jusqu'à la mort de l'animal avec une excrétion de virus maximale entre 3 et 5 semaines post infection (Witter *et al.*, 1971). Plus précisément, après desquamation, les particules virales associées aux débris de phanères peuvent réinfecter d'autres poulets lorsqu'elles sont inhalées par ces derniers. Il faut noter que le virus ainsi associé aux débris de phanères, peut rester infectieux pendant plusieurs mois entre 20-25°C, jusqu'à plusieurs années à 4°C.

3.1.3. Les formes cliniques

La pathologie associée à l'infection par GaHV-2 est complexe et les syndromes cliniques et pathologiques varient en fonction de nombreux paramètres influençant la mortalité et la gravité des lésions, parmi lesquels, pour ne citer que les plus importants :

- La virulence de la souche virale ; en effet, les souches moyennement oncogènes (v, pour virulente) causent généralement des lésions des nerfs périphériques. Par contre, les pathotypes les plus virulents (vv et vv+, pour very virulent et very virulent +) apparus récemment suite à des épisodes d'échappement à la vaccination, induisent de plus nombreux lymphomes viscéraux associés à une mortalité plus importante.

- La génétique de l'hôte ; le complexe majeur d'histocompatibilité ou complexe B localisé sur le microchromosome 16 du poulet joue un rôle majeur dans la résistance à la maladie de Marek. Ainsi des poulets âgés de 21 jours homozygotes B^{13}/B^{13} développent des lymphomes à la suite d'une infection par des souches virulentes ou hypervirulentes GaHV-2 contrairement aux poulets homozygotes B^{21}/B^{21} . Ces derniers présentent une virémie plus faible que les poulets sensibles pouvant expliquer l'absence de développement des lymphomes.

- Le sexe de l'hôte ; les femelles semblent plus sensibles à la maladie que les mâles.

- L'âge au moment de la première exposition ; les jeunes oiseaux semblent plus sensibles à la pathologie (Calnek, 2001).

- Les voies d'inoculation ; expérimentalement les voies d'inoculation intrapéritonéale, intratrachéale ou intramusculaire conduisent plus rapidement à un développement des lymphomes que la contamination aérienne classique (Debba-Pavard *et al.*, 2008a; Fabricant *et al.*, 1977; Islam *et al.*, 2001).

- Le mode d'élevage ; certaines conditions d'élevage stressantes ont également une incidence sur le développement de la maladie de Marek.

Au sein des élevages, la maladie de Marek s'exprime principalement sous deux formes : la forme neuronale décrite par J. Marek dite classique, et la forme lymphomateuse dite aiguë.

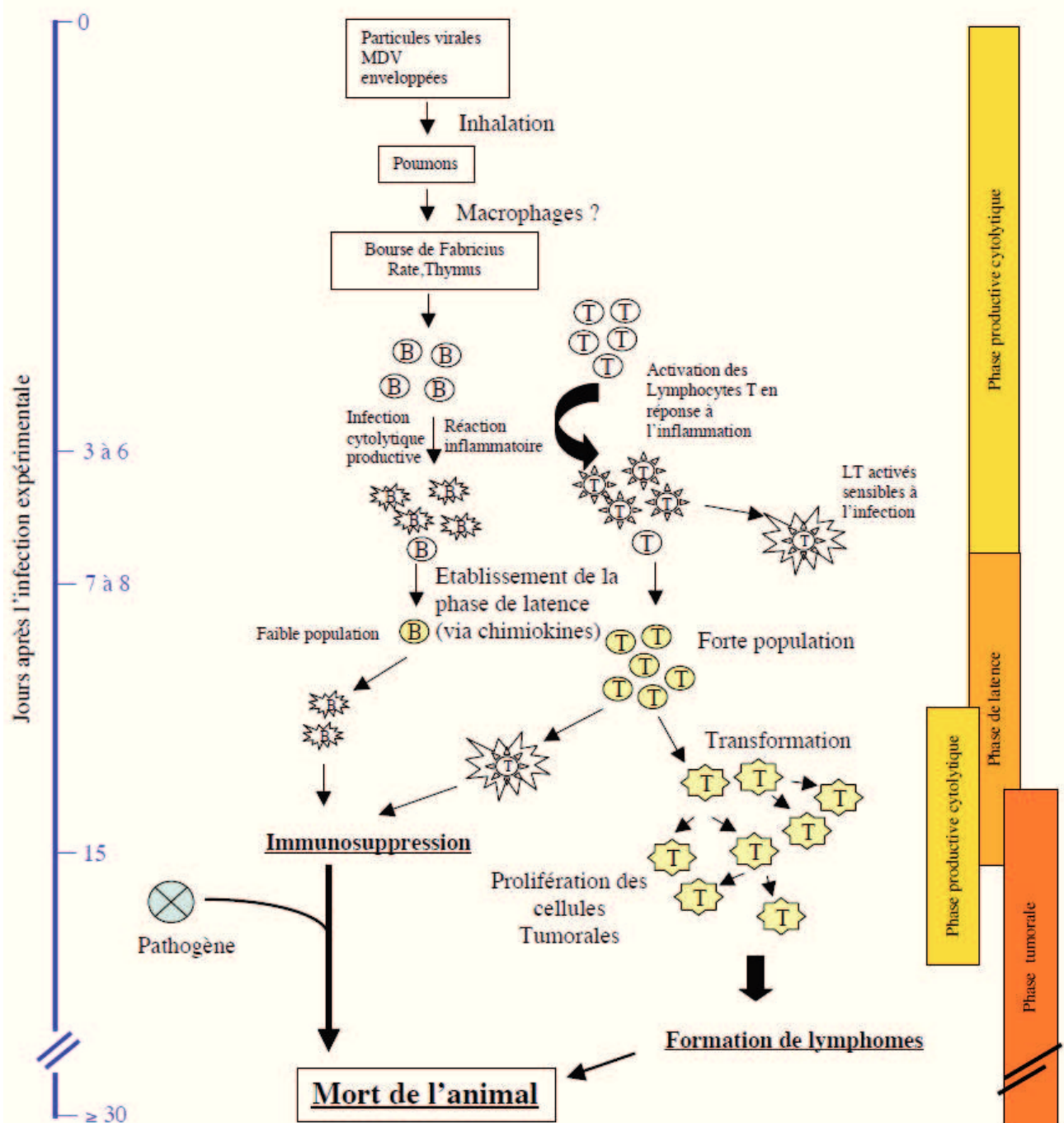


Figure 24 : Chronologie de la pathogénèse d'une souche oncogénique de GaHV-2. Les différentes étapes depuis la pénétration du virus par les voies respiratoires de l'animal jusqu'à sa mort sont décrites. Les cellules intervenant dans le cycle viral et leur état d'activation sont représentées ainsi:

- | | | | | | | | | | |
|-----|---------------------------------|----------------------|---|-----|-----------------------|----------------------|-----------------------|--------------------------------|---------------------------|
| (B) | Lymphocytes B normaux | Cellule B avec virus | Infection cytotytique des cellules B | (T) | Lymphocytes T normaux | Cellule T avec virus | Lymphocytes T activés | Cellule T avec virus et étoile | Lymphocytes T transformés |
| (B) | Infection latente de cellules B | | | | | | | | |
| (T) | Infection latente de cellules T | Cellule T avec virus | Infection lytique des lymphocytes T activés | | | | | | |

La forme classique de la maladie débute lentement après une longue période d'incubation (3 à 9 semaines) et se manifeste dans un premier temps par des troubles locomoteurs. Quelques jours plus tard, l'apparition d'une boiterie généralement unilatérale se manifeste, pouvant être accompagnée de divers signes de paralysie traduisant l'infiltration des nerfs périphériques par les cellules lymphoïdes : ailes pendantes, difficultés respiratoires, rétention fécales... Une atteinte oculaire peut également se manifester par un changement de la forme et de la couleur de l'iris, ainsi que de la forme et de la position de la paupière, pouvant aboutir à une cécité partielle ou totale de l'animal. La mort survient généralement lorsque la paralysie engendre une incapacité des animaux à accéder à la nourriture.

La forme aigüe est caractérisée par une évolution plus rapide de la maladie, aboutissant à la mort de l'animal en environ 4 semaines. L'animal se retrouve abattu, amaigri, anémique avec également une paralysie progressive. Cependant cette forme de la maladie se caractérise surtout par l'apparition de tumeurs multiples pouvant être observées au niveau des organes et tissus tels les poumons, le proventricule, les reins, le foie, le cœur, le pancréas, l'intestin, les ovaires, les testicules, le mésentère, les sacs aériens et la peau où des nodules tumoraux apparaissent à la base des plumes. La fréquence et la sévérité de ces tumeurs sont liées à la fois à la souche virale infectante et à la lignée génétique de l'animal.

3.1.4. Physiopathologie virale

L'infection d'un poulet *in vivo* par une souche oncogénique du virus de la maladie de Marek se présente généralement sous quatre phases (Figure 24):

- Une infection cytotytique précoce.
- Une infection latente.
- Une seconde phase cytotytique dite tardive.
- Une phase proliférative mettant en jeu les cellules lymphoïdes infectées conduisant à la formation de lymphomes.

L'infection cytotytique précoce se produit après l'entrée du virus par les voies respiratoires. Le virus sous forme libre semble être phagocyté par les macrophages ou par les cellules dendritiques pulmonaires (Barrow *et al.*, 2003). Ces cellules immunitaires infectées véhiculent le virus GaHV-2 vers les organes lymphoïdes qui lors d'inoculation expérimentale intra trachéale se retrouvent atteints dès 24 à 36 heures pi (post-infection) (Baigent & Nair, 2004; Schat & Nair, 2008b). Au sein des organes lymphoïdes secondaires et surtout de la rate, une première infection productive cytotytique se met en place entre 3 et 6 jours pi (Shek *et al.*, 1983). Elle se caractérise par une dissémination du virus de cellule à cellule, sans production de virions libres. Elle touche principalement les lymphocytes B et dans une moindre mesure les lymphocytes T et les macrophages (Barrow *et al.*, 2003; Shek *et al.*,

1983). La lyse de ces cellules induisant une inflammation entraîne le recrutement des macrophages, des lymphocytes T et B ainsi que des granulocytes. Différentes cytokines sont surexprimées durant l'inflammation (IFN- γ , IL-1 β , IL-8, IL-6, IL-18), avec également une activation de la voie de synthèse de l'oxyde nitrique NO (iNOS : inducible NO synthase, NO : Nitric Oxide). L'activation des lymphocytes T consécutive à cette réponse inflammatoire faciliterait leur infection par le virus GaHV-2.

Après 6 à 7 jours, une phase d'infection latente s'installe principalement au sein des lymphocytes T CD4⁺ activés et d'une manière moins importante au sein des lymphocytes T CD8⁺ et des lymphocytes B (Baigent & Nair, 2004; Shek *et al.*, 1983). L'établissement et le maintien de la latence est un processus multifactoriel impliquant des gènes viraux (le gène *LAT*, le gène *meq* ou certains miARN...) mais également différents facteurs cellulaires. Notamment, le facteur de maintien en latence (LMF : latency maintainig factor), l'oxyde nitrique NO, l'IFN- α et l'IFN- γ , (Schat & Nair, 2008b; Xing & Schat, 2000), ces derniers semblant bloquer la réplication virale avant la traduction des gènes tardifs (Levy *et al.*, 2003a).

Durant la phase d'infection latente, il est observé un phénomène de réactivation pour quelques cellules en latence, aboutissant à une seconde phase d'infection cytolytique dite tardive, à partir de 2 à 3 semaines pi. Le développement et l'étendue de cette seconde infection cytolytique est dépendante de la résistance génétique des hôtes et de la virulence de la souche virale. Elle se manifeste au niveau des organes lymphoïdes ainsi que dans certains tissus (rein, pancréas, proventricule). Aucune particule virale libre n'est détectable dans le sang périphérique, mais la forte association du virus aux cellules indiquerait une dissémination virale de cellule à cellule. Par ailleurs, dès la 3^{ème} semaine, le GaHV-2 se réplique essentiellement dans les kératinocytes des follicules plumeux. Il faut noter que ces cellules sont l'unique site de réplication productive de particules virales libres, à l'origine de la transmission horizontale du virus (Calnek *et al.*, 1970).

La dernière phase de la maladie ou phase tumorale spécifique des souches oncogènes de type GaHV-2 apparaît classiquement lors d'une contamination expérimentale à partir de la troisième semaine suivant l'infection. Cette phase se caractérise par le développement de lymphomes T issus dans la majorité des cas de lymphocytes T CD4⁺ infectés de manière latente. Par ailleurs, ces cellules expriment fortement le marqueur CD30 ce qui fait de la maladie de Marek un bon modèle de lymphome Hodgkinien. Cette dernière phase est également accompagnée d'une immunosuppression permanente affectant à la fois l'immunité humorale et l'immunité à médiation cellulaire, pouvant entraîner une infection létale par un autre pathogène. En conclusion, la lymphomagenèse dépend de six critères :

- La susceptibilité de l'hôte à l'infection.
- Un contrôle intrinsèque et extrinsèque de la réplication virale.

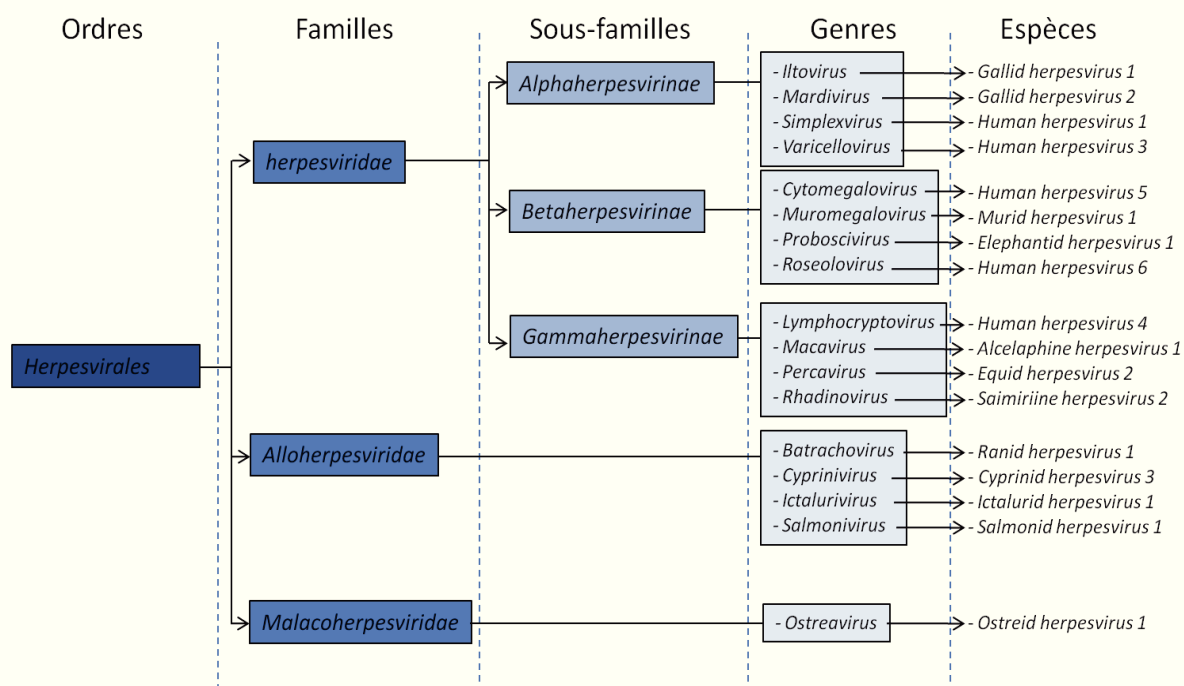


Tableau 1 : Classification de l'ordre des *Herpesvirales*. L'espèce type de chaque genre est représentée.

- Une division cellulaire avec le génome viral intégré.
- L'expression d'oncogènes viraux tels que le gène *meq* ou les miARN.
- L'activation d'oncogènes cellulaires.
- La répression de l'induction de l'apoptose.

3.2. Le virus de la maladie de Marek

3.2.1. Classification

Le virus GaHV-2 (Gallid Herpesvirus de type 2) fait partie de l'ordre des *Herpesvirales* (Tableau 1), de la famille des *Herpesviridae*, de la sous famille des *Alpha-herpesvirinae*, et du genre *Mardivirus*. L'ordre des *Herpesvirales* se divise en trois familles (Tableau 1). La famille des *Herpesviridae* comportant exclusivement des virus de mammifères et d'oiseaux, la famille des *Alloherpesviridae* qui se compose de virus d'amphibiens et de poissons, et la famille des *Malacoherpesviridae* ne comportant qu'une seule espèce infectant l'huître.

A l'origine la famille des *Herpesviridae* a été divisée en trois sous famille en fonction des propriétés biologiques des virus. Ainsi la sous-famille des *Alpha-herpesvirinae* a été définie sur la base d'un cycle réplcatif court, d'une dissémination rapide en culture de cellules, d'une forte destruction des cellules infectées et de l'établissement principalement mais pas exclusivement de leurs infections latentes au sein des ganglions sensoriels. Cette sous-famille comprend les virus humains herpes simplex virus 1 et 2 (HSV-1 ou HHV-1 et HSV-2 ou HHV-2), ainsi que le virus de la varicelle et du zona (VZV ou HHV-3). La sous famille des *Beta-herpesvirinae* présente un cycle réplcatif long, une infection lente en culture de cellules. Elle comprend les virus humains HHV-6, HHV-7 ainsi que le cytomégalo virus (HCMV ou HHV-5). La dernière sous-famille, les *Gamma-herpesvirinae* se répliquent exclusivement dans des cellules lymphoblastoïdes, infectent de manière lytique les cellules épithéliales et les fibroblastes et se retrouvent fréquemment de manière latente au sein des tissus lymphoïdes. Parmi cette sous famille, sont retrouvés le virus d'Epstein-Barr (EBV ou HHV4) ainsi que le virus de Kaposi (KSHV : Kaposi Sarcoma Herpes Virus ou HHV8).

Du fait de ses propriétés biologiques et plus particulièrement de sa capacité à induire des lymphomes T ainsi que de son faible potentiel réplcatif en culture cellulaire, GaHV-2 a longtemps été associé au virus EBV et classé parmi les *Gamma-herpesvirinae*. Cependant, l'étude du génome de GaHV-2 a conduit à son reclassement dans la sous-famille des *Alpha-herpesvirinae*, de par son organisation génomique et plus précisément de la présence de structures répétées caractéristiques de cette sous-famille.

Génotype	Pathotype	Oncogénicité	Souches
Gallid herpesvirus type 2 (GaHV2)	Hyper hypervirulent (vv+)	+++	RK1, 584A, 648A
	Hypervirulent (vv)	++	RB-1B, Md-5, Md-11
	Virulent (v)	+	GA, HPRS16, JM
	Faiblement virulent (m)	nulle	Rispens, Cu-2, HPRS17
Gallid herpesvirus 3 (GaHV3)	Faiblement virulent	nulle	SB-1, HPRS24, HN-1
Meleagrid herpesvirus type 1 (MeHV-1)	Faiblement virulent	nulle	HVT, WTHV, HPRS26

Tableau 2 : Classification du genre *Mardivirus*.

Au sein du genre *Mardivirus*, trois génotypes ont été décrits (Tableau 2). Le génotype 1 *Gallid herpesvirus 2* (GaHV-2 anciennement nommé MDV-1) regroupant la totalité des souches oncogéniques (RB-1B, md5...), faiblement oncogéniques (HPRS16, GA...) ou naturellement non oncogéniques telle la souche vaccinale Rispens CVI988 (Rispens *et al.*, 1972a; Rispens *et al.*, 1972b), ainsi que les souches atténuées par passage en cultures cellulaires issues des souches mères HPRS16 et Rispens CVI988. Le génotype 2 correspond à l'espèce *Gallid herpesvirus 3* (GaHV-3 anciennement nommé MDV-2) qui comporte des souches non pathogènes et non oncogènes, dont certaines ont été utilisées comme vaccin (souche SB-1) (Witter, 1987). Le génotype 3, dénommé *Meleagrid herpesvirus 1* (MeHV-1 anciennement Herpes Virus de Turkey : HVT) rassemble les souches non oncogènes du dindon qui ont également été très utilisées en tant que souches vaccinales depuis les années 1970 (Witter *et al.*, 1970).

3.2.2. Vaccin

Compte-tenu des pertes économiques potentielles engendrées par GaHV-2, tous les poulets d'élevage de la planète sont vaccinés contre la maladie de Marek. En effet, les pertes peuvent atteindre près de 40% des poulets non vaccinés, et aucun traitement curatif n'existe à ce jour. Seule une prophylaxie vaccinale dans les premiers jours de la vie de l'animal est efficace. Les souches vaccinales les plus couramment utilisées sont des souches non pathogènes de GaHV-2 (CVI988) (Rispens *et al.*, 1972a; Rispens *et al.*, 1972b), des souches de GaHV-3 (SB-1, 301B/1) (Witter, 1987) ou des souches de MeHV-1 (Fc126, HPRS24). Ce sont des vaccins vivants utilisés soit individuellement soit en combinaison. Historiquement, le premier vaccin était constitué de la souche de GaHV-2 HPRS16 atténuée par passages successifs en culture de cellules primaires de canard (Churchill *et al.*, 1969). Rapidement après sa découverte en 1970, la souche vaccinale HVT (Witter *et al.*, 1970) fut largement utilisée à travers le monde, et elle demeure à ce jour le plus souvent protectrice bien que des souches d'échappement apparaissent. Au milieu des années 1980, dans le but de tenter de résoudre les problèmes d'échappement aux souches vaccinales, les producteurs de vaccins ont généré des vaccins bivalents, en combinant HVT avec des souches SB-1 ou 301B/1 aux Etats-Unis ou avec la souche CVI988 en Europe. Depuis des vaccins trivalents regroupant CVI988, HVT et SB-1 ont vu le jour afin de répondre à l'apparition de souches de plus en plus virulentes. Des voies alternatives ont été explorées pour la création de vaccins recombinants : a) l'utilisation d'un virus de la variole aviaire (fowpox) recombinant exprimant le gène de la glycoprotéine gB de GaHV-2 (Omar & Schat, 1996), b) l'utilisation de souches de GaHV-2 génétiquement atténuées, dont le génome a été délété d'un ou plusieurs gènes. Notamment, l'équipe de Witter a créé la souche vaccinale RM1, qui avait perdu son pouvoir oncogène tout en gardant sa capacité à se répliquer *in vivo*, en insérant des LTR (Long

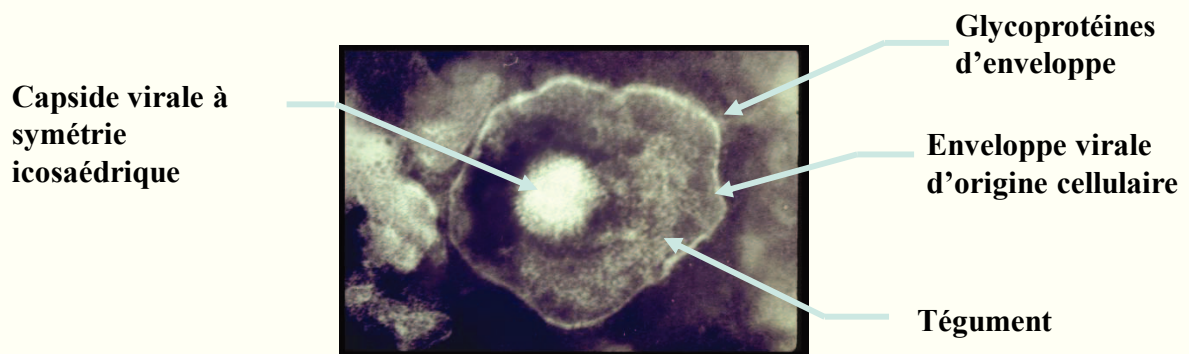


Figure 25 : Micrographie électronique du virion MDV. (Cliché F. Coudert)

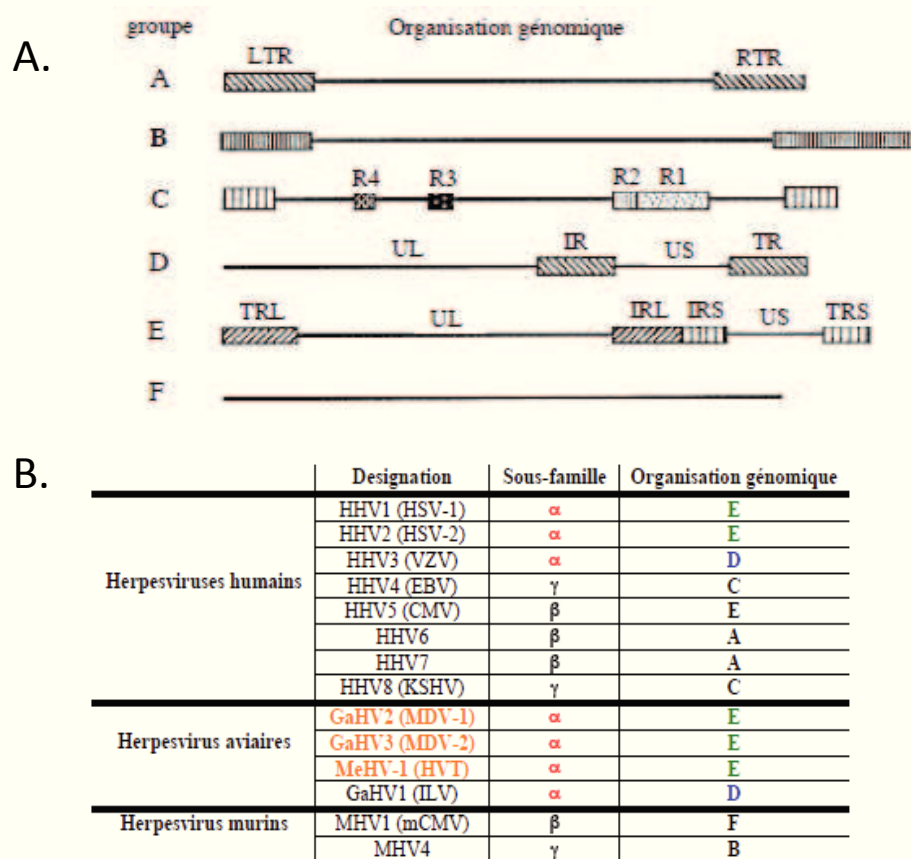


Figure 26 : Organisation génomique des herpesvirus des groupes A à F (Pellet and Roizman., 2007).

(A) Représentation schématique des six classes de génomes retrouvées chez les virus de la famille des *Herpesviridae*. Les régions uniques sont représentées par les lignes horizontales et les régions répétées sont figurées par des rectangles et sont désignées : domaine répété gauche et droit (LTR et RTR) pour le groupe A, domaines répétés internes R1 à R4 pour le groupe C et domaines répétés internes et terminaux (IR et TR) pour le groupe D. Dans le groupe B, les séquences terminales sont répétées en nombre variable à chaque extrémité et entre les deux extrémités du génome. Le groupe E est composé d'une séquence unique longue (U_L) et d'une séquence unique courte (U_S) toutes deux encadrées par des séquences répétées internes et externes (TR_L , IR_L , IR_S et TR_S). Aucune répétition terminale n'a été décrite pour le sous groupe F. (B) Classification d'herpesvirus humains, aviaires et murins et organisation génomique correspondante.

Terminal Repeat) du rétrovirus REV (Reticuloendotheliosis Virus) à l'intérieur du génome de la souche virulente JM (Witter *et al.*, 1997). Plus récemment, des essais utilisant des bacmides recombinants délétés pour l'oncogène *meq* ont abouti à une protection contre des souches hyper-virulentes vv+, supérieure à celle obtenue avec la souche CVI988 et comparable voire supérieure à celle obtenue avec les vaccins bivalents, en conditions expérimentales (Lee *et al.*, 2010; Lee *et al.*, 2008).

3.2.3. Structure

Le virus de la maladie de Marek présente une structure typique pour un herpesvirus (Figure 25). Les particules virales observées dans des inclusions cytoplasmiques de follicules plumeux présentent un diamètre compris entre 270 et 400 nm, alors que les particules observées dans les fins espaces intercellulaires présentent un diamètre plus faible de l'ordre de 150 nm (Calnek *et al.*, 1970). Le virion est constitué d'une capside icosaédrique, mesurant de 85 à 100 nm de diamètre et composée de 162 capsomères, renfermant l'ADN viral. Cette capside est entourée d'un amas organisé de protéines correspondant au tégument, qui assure la continuité avec les glycoprotéines implantées dans l'enveloppe d'origine cellulaire.

3.2.4. Organisation génomique de GaHV-2

Les génomes viraux des trois sérotypes du virus de la maladie de Marek sont composés d'un ADN linéaire bicaténaire d'environ 160 à 180 kpb. Ils possèdent une organisation de classe E identique à celle de l'herpesvirus simplex de type 1 (HHV-1) (Figure 26). Il faut noter que cette organisation de type E est majoritaire au sein de la sous-famille des *Alpha-herpesvirinae* mais n'est pas exclusive puisque des génomes de type D sont également observés pour certains virus de la famille tel que VZV. Les génomes de classe E se composent d'une sous-unité longue U_L et d'une sous-unité courte U_S toutes deux encadrées de régions répétées inversées identiques TR_L/IR_L (Terminal Repeat Long et Internal Repeat Long) pour U_L et IR_S/TR_S (Internal Repeat Short et Terminal Repeat Short) pour U_S . De plus, des séquences télomériques composées de répétitions d'hexamères GGTTAG de longueur variable sont retrouvées au niveau de la jonction IR_L/IR_S ainsi qu'aux extrémités du génome (Kishi *et al.*, 1991). Ces répétitions télomériques pourraient entre autre être impliquées lors de la phase de latence dans l'intégration du génome viral, qui s'effectue principalement au niveau des télomères des chromosomes de l'hôte (Delecluse & Hammerschmidt, 1993). En effet, contrairement aux autres herpes virus dont le génome est essentiellement présent sous une forme épisomale durant les phases de latence, le génome du GaHV-2 est retrouvé intégré aux génomes de l'hôte.

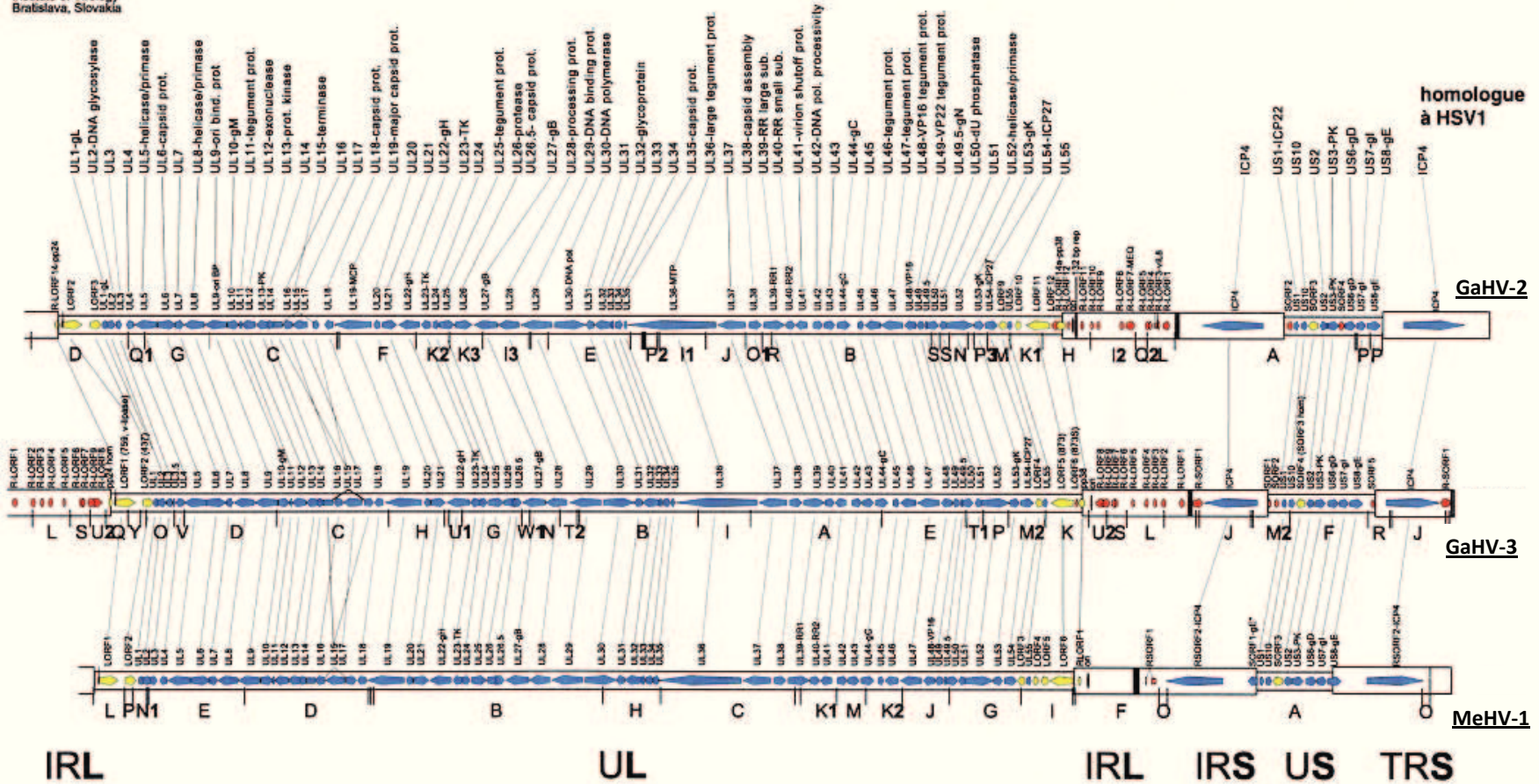


Figure 27 : Comparaison de la structure génomique des 3 génotypes de Mardivirus, GaHV-2, GaHV-3 et MeHV-1. Les gènes colorés en bleu ont des homologues chez les autres alphaherpesvirus. Les gènes en jaune sont spécifiques aux trois génotypes de MDV et les gènes en rouge sont spécifiques de l'espèce. Abréviations : hom., homologue ; pol., polymérase ; prot., protéine ; rep., repeat ; RR, ribonucleotide reductase ; sub., subunit ; TK, thymidine kinase.

Au début des années 2000, les génomes complets de souches appartenant aux trois espèces du genre *Mardivirus* ont été séquencés, notamment le génome des souches GaHV-2-GA (AF147806) et GaHV-2-Md5 (AF243438), de la souche GaHV-3-HPRS24 (NC_002577), et de la souche MeHV-1-FC126 (NC_002641) (Afonso *et al.*, 2001; Izumiya *et al.*, 2001; Lee *et al.*, 2000; Tulman *et al.*, 2000). Bien que leurs tailles et leurs compositions en GC soient variables, la comparaison des séquences des trois espèces de *Mardivirus* a permis de mettre en évidence une organisation génique globalement similaire (Figure 27). Ainsi, la souche GA possède un génome de 174040pb dont 43,9% de GC, Md5 de 177874pb dont 44,1% de GC, HPRS24 de 164270pb dont 53,6% de GC et FC126 de 159160pb dont 47,6% de GC. Le nombre de cadres ouverts de lecture ou ORF (Open Reading Frame) entre les génomes de MeHV-1, GaHV-2 et GaHV-3 est également très proche, avec environ une centaine d'ORF retrouvés au sein des trois espèces, et les différences majeures retrouvées sont localisées au niveau des séquences répétées TR_S/IR_S et TR_L/IR_L.

3.2.4.1. Origine de réplication lytique

Les trois sérotypes du genre *Mardivirus* comportent des origines de réplication lytique (Ori_{lyt}) dans leurs régions répétées TR_L et IR_L. Comme pour les autres *alphaherpesvirus*, elles se caractérisent par une séquence riche en AT (deux répétitions directe de TATATTATAT), flanquée de sites putatifs de liaison avec la protéine OBP (Origin Binding Protein) codée par le gène U_L9 (Bradley *et al.*, 1989a; Camp *et al.*, 1991; Katsumata *et al.*, 1998). Ces sites de fixation sont caractérisés par une séquence de 9 pb très conservée (CGTTTCGCAC). Cette structure spécifique, en association avec différentes protéines précoces, va permettre l'initiation de la réplication de l'ADN viral en mode de cercle roulant (ou rolling circle).

Outre le rôle d'origine de la réplication de l'ADN viral, cette région possède un rôle de promoteur bidirectionnel permettant l'expression des gènes situés de part et d'autre de l'Ori_{lyt} (Figure 28). A droite de cette dernière, vers les régions répétées TR_L/IR_L, sont transcrits les gènes potentiellement impliqués dans la latence et la transformation (famille des gènes 1,8kb, *meq*, *vIL-8*...) tandis que, à gauche, vers la région U_L, s'effectue la transcription des gènes codant pour pp38/pp24 qui sont associées à la phase précoce de l'infection lytique et au maintien de la transformation (Figure 28). Ce promoteur a été relativement bien décrit et présente de nombreux éléments *cis*-régulateurs (Figure 28) (Shigekane *et al.*, 1999). Parmi ces éléments, deux sites de fixation MERE II pour des homodimères Meq/Meq permettent d'induire une diminution de l'expression de pp38 et des gènes de 1,8kb (Levy *et al.*, 2003b). De plus, il a été montré qu'un hétéropolymère constitué de pp38 et de pp24 se fixerait au niveau du promoteur sur la séquence CTGCTCATTT. Il permettrait d'augmenter l'efficacité transcriptionnelle dans les deux sens (Ding *et al.*, 2008; Ding *et al.*, 2007; Ding *et al.*, 2006). L'expression dans un sens ou dans l'autre du promoteur est inégale et une activité 5 à 15 fois

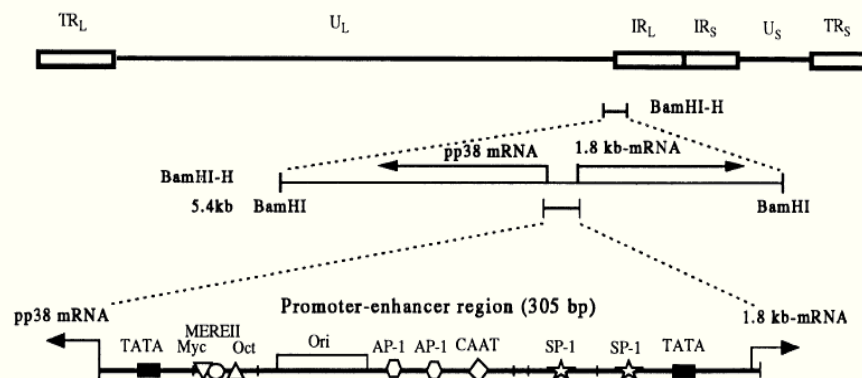


Figure 28 : Description et localisation du promoteur bi-directionnel présent dans la région BamHI-H du génome de GaHV-2 (Shigekane et al., 1999). Parmi les élément cis-régulateur potentiels, sont représentés, les boîtes TATA (rectangle noir), le site Oct-1 (triangle), la boîte CAAT (losange), le site Myc (triangle inversé), les sites SP-1 (étoiles), le site de reconnaissance de Meq : MERE II (cercle) et les sites AP1 (hexagone). Les sites d'initiations de la transcription du gène pp38 et des gènes de la famille des 1,8kb sont indiqués par des flèches.

supérieure a été retrouvée au niveau du promoteur de la famille des gènes 1,8kb en comparaison du promoteur de pp38 lors d'infection lytique de différentes souches de GaHV-2 en fibroblastes embryonnaires de poulets (CEF) (Chen *et al.*, 2009; Shigekane *et al.*, 1999). Cette région promotrice est complexe, et des études complémentaires sont nécessaires afin de comprendre et de définir l'expression des gènes de cette région durant la latence, la réactivation et l'infection lytique.

3.2.4.2. Variabilité génique au sein des isolats de GaHV2

La comparaison des séquences des différents génomes des souches de GaHV-2 fait apparaître une très forte similarité entre elles. Outre des variations au sein de certaines ORF (U_L36 et U_L49), les différences correspondent essentiellement à un fort polymorphisme au niveau des jonctions TR_S/U_S, IR_S/U_S et de la région télomérique entre TR_L et TR_S (Spatz & Silva, 2007b). En outre, un nombre variable de répétitions des séquences dites de 132 pb et 60 pb présentes respectivement dans l'IR_L/TR_L et l'IR_S/TR_S ont été mises en évidence (Spatz & Silva, 2007b). De même au sein de l'IR_L/TR_L, les séquences des gènes *RLORF12*, *RLORF6*, *23kD*, *meq* et *RLORF4* semblent varier entre les différentes souches virales. Cependant, il faut noter que ces mêmes variations ne semblent pas correspondre à des signatures de souches car elles peuvent également apparaître au sein d'une même souche lors d'infection *in vivo* successives ou *in vitro* lors de passages successifs en culture de cellules.

3.2.4.2.1. Répétitions des 132 pb

L'implication dans la latence et l'oncogénicité des répétitions de 132 pb a longtemps été suspectée. En effet, sur la base d'observations relatant la corrélation apparente entre l'augmentation du nombre de séquences de 132 pb et le nombre des passages lors de l'atténuation des souches oncogéniques en culture de cellules, il avait été suspecté un rôle potentiel de cette région dans l'oncogénèse induite par GaHV-2 (Bradley *et al.*, 1989b). D'une manière générale, les souches virulentes de GaHV-2 possèdent seulement deux copies de ces répétitions de 132 pb, mais de multiples copies pouvant aller jusqu'à plus d'une trentaine peuvent être retrouvées dans les génomes après passage des virus en culture de cellules. Cependant, les résultats récents semblent contredire ces hypothèses de lien entre l'oncogénicité d'une souche et le nombre de copies de séquence de 132 pb au sein de son génome. A titre d'exemple, la souche hyper virulente RB-1B contient neuf copies de ces répétitions (Spatz *et al.*, 2007), tandis que la souche vaccinale CVI-988 n'en contient que deux (van Iddekinge *et al.*, 1999). De plus, grâce à l'utilisation de virus recombinants possédant aucune, deux ou neuf copies de cette séquence, Silva et ses collaborateurs ont montré que ces répétitions n'étaient pas essentielles à la pathogénèse induite par GaHV-2 (Silva *et al.*, 2004). Par contre, un lien semblerait exister entre le nombre de copies de cette séquence et l'efficacité des souches à se répliquer en culture

de cellules (Silva *et al.*, 2004). D'autant qu'une pression de sélection inverse a été observée sur ces répétitions *in vivo*. En effet, après plusieurs infections successives de poulets par une souche atténuée contenant de multiples copies des 132 pb au sein de son génome, les auteurs ont montré que rapidement le nombre de copies tombait à deux ou trois (Silva *et al.*, 2004). Enfin, une étude plus récente a mis en évidence que l'atténuation de la souche virale par passages en culture de cellules n'était pas due au nombre de copies des 132 pb car un effet identique d'atténuation de la souche a été observé entre un virus délété des 132 pb et le virus sauvage (Silva & Gimeno, 2007). Il semblerait donc que la variation du nombre de copies de cette séquence soit plutôt une conséquence du passage du virus en culture de cellules sans impact réel sur sa virulence.

3.2.4.2.2. Répétitions de 60 pb

Le nombre des répétitions de 60 pb est également variable selon les souches (Spatz & Silva, 2007b), et une fois encore, un lien a été proposé entre la présence de ces répétitions et le pouvoir oncogène des souches. En effet, ces séquences, non présentes au sein des génomes de la souche vaccinale Rispens CVI988 ou de la souche moyennement virulente CU-2, sont retrouvées au nombre de 2 pour la souche GA, 7 pour Md5, 11 pour RB-1B et peuvent aller jusqu'à plus d'une vingtaine de répétitions dans certaines souches. Par ailleurs, notre équipe TLVI a montré récemment que ces répétitions constituaient en réalité le promoteur du gène *LAT* ainsi que des miARN du cluster 8-10, introniques au sein de cet ARN. Chaque unité de 60 pb renferme un élément de réponse fonctionnel pour la protéine cellulaire p53 (Stik *et al.*, 2010). De plus, deux répétitions semblent nécessaires et suffisantes pour une activité promotrice optimale, la multiplicité de ces répétitions ne semblant pas modifier de façon significative cette activité (Stik *et al.*, 2010).

3.2.5. Transcriptome, protéome et cycle viral de GaHV-2

Une fonction réelle (démontrée biologiquement) ou putative (par analogie de séquence avec les autres herpesvirus) a été attribuée à la grande majorité des gènes des régions U_L et U_S . En effet, ces deux régions uniques semblent renfermer les gènes viraux « core » colinéaires avec ceux des autres herpesvirus (Figure 27). Ainsi, la région U_L de chacun des trois sérotypes de GaHV-2 renferme les gènes U_L1 à U_L55 homologues à HHV-1. Ces gènes possèdent un pourcentage d'identité variable entre HHV-1 et GaHV-2, cependant la colinéarité de ces gènes étant particulièrement forte cela a permis par analogie, de suspecter leurs implications dans la réplication de l'ADN et dans la morphogenèse de GaHV-2 (glycoprotéines, protéines de capsid et de tégument). La région U_S des trois sérotypes de

GaHV-2 encode pour les gènes U_{S1} , U_{S2} , U_{S3} , U_{S6} , U_{S7} , U_{S8} et U_{S10} homologues à HHV-1 ; cependant il faut noter que pour la région U_S les gènes homologues à U_{S4} , U_{S5} , U_{S9} , U_{S11} , U_{S12} ne sont pas présents dans le génome des trois *Mardivirus*.

Par ailleurs, l'existence de plusieurs gènes spécifiques des *Mardivirus*, localisés dans les régions uniques U_L et U_S , permettrait de leur supposer une fonction conservée pour la mise en place du cycle dans les espèces aviaires. Ainsi, huit ORF ont été identifiées : LORF2, LORF3, LORF9, LORF10, LORF11, R-LORF14, R-LORF14a et SORF3. Hormis cette dernière, ces gènes sont retrouvés proches des jonctions TR_L/U_L et U_L/IR_L au voisinage des gènes « core » précédents. En effet, la localisation des gènes spécifiques d'espèce virale aux extrémités des gènes « core » est classique chez les grands virus à ADN, et elle témoignerait d'une acquisition récente liée à l'adaptation à l'hôte. A ce jour, seuls trois de ces gènes ont été investigués. Une fonctionnalité a été attribuée à LORF2, R-LORF14a et R-LORF-14 codant respectivement pour une lipase virale (retrouvée chez les trois génotypes), pp38 et pp24 (retrouvées chez GaHV-2 et GaHV-3).

Enfin, le dernier type de gènes spécifiques à chaque génotype de GaHV-2 correspond aux gènes « satellites » localisés principalement dans les régions répétées inversées du virus ou dans les régions proches de leurs jonctions. Dans ce type de gènes sont notamment retrouvés les gènes impliqués dans l'oncogenèse des souches GaHV-2, et hormis *ICP4*, les gènes présents dans ces régions répétées sont généralement spécifiques de chaque génotype.

3.2.5.1. Le cycle lytique

En corrélation avec la chronologie du cycle des herpesvirus déterminée en culture cellulaire *in vitro*, dans des conditions d'infection lytique, les gènes ont été répartis en trois groupes (Tableau 3) (Schat *et al.*, 1989) :

- les gènes α également appelés gènes très précoces ou IE (Immediate Early) correspondant à des transactivateurs des gènes viraux et cellulaires.
- Les gènes β également appelés gènes précoces ou E (Early) sont transactivés par les gènes α . Ce sont tous les gènes impliqués dans la réplication du génome viral (facteurs de transcription, protéines du complexe polymérasique, protéine se liant à l'origine de réplication lytique, enzymes virales impliquées dans le métabolisme des acides nucléiques).
- Les gènes γ également appelés gènes tardifs ou L (Late) sont transactivés en fin de cycle par les gènes α suite à leur démasquage lors de la réplication du génome durant la phase précoce. Ils correspondent aux protéines de structure, tels que les protéines de la capside, du tegument ou les glycoprotéines d'enveloppe.

	protéines	Fonction
Protéines très précoces (α)	ICP4 ICP27	Protéines transactivatrices des gènes précoces et tardifs
	gK	Glycoprotéine impliquées dans la diffusion virale de cellule à cellule
Protéines précoces (β)	pp24 pp38 Thymidine Kinase DUTPase Ribosyl reductase ADN polymérase Hélicase-Primase Protéine de liaison à l'ADN simple brin	Facteurs de transcription Protéines impliquées dans la réplication de l'ADN et le métabolisme des acides nucléiques
Protéines tardives (γ)	gB gC gD gH gL	Glycoprotéines impliquées dans l'entrée du virus dans les cellules
	gE/gI gM/gN	Glycoprotéines impliquées dans la diffusion virale de cellule à cellule
	VP11/12 VP22 VP13/14	Protéine de tégment Protéine de tégment Protéine de tégment
	VP16	Protéine de tégment transactivatrice des gènes très précoces
	VP5 VP19c VP21-VP24 pré-VP22a VP23 VP26	Protéine majeure de capsid Protéine de capsid Protéines de capsid Protéine d'échafaudage Protéine de capsid Protéine de capsid

Tableau 3 : Les principales protéines du MDV.

La régulation de ces gènes se fait en cascade, et dès lors qu'un groupe de gènes est activé, il va à son tour activer le groupe suivant de façon coordonnée et avec souvent un rétrocontrôle négatif sur le ou les groupes précédents.

Bien que les protéines servant au cycle réplcatif de GaHV-2 ne soient que peu caractérisées, il est admis du fait de l'homologie et de la colinéarité des gènes « core » une régulation du cycle lytique commune entre les herpesvirus. Ainsi, HHV-1 qui est le plus extensivement étudié sert de modèle pour l'étude de la réplication de la famille des herpesvirus.

3.2.5.1.1. Entrée du virus.

L'entrée du virus consiste en la fixation du virion à la membrane cellulaire, puis à la fusion de cette dernière avec la membrane du virion permettant l'entrée de la capside et des protéines de tégument au sein du cytoplasme. Cette entrée se réalise en trois étapes. La première met en jeu les glycoprotéines gC et gB présentes sur l'enveloppe virale. En se fixant sur les glycosaminoglycanes exposés à la surface des cellules, elles vont entraîner un changement conformationnel du virus permettant l'initiation de la fusion des membranes. Cette seconde étape est induite par la liaison de la glycoprotéine gD aux récepteurs cellulaires de type nectines, HVEM (Herpes Virus Entry mediator) de la famille des TNF (Tumor Necrosis Factor) ou de type héparane sulfate (Roizman *et al.*, 2007). Bien que le mécanisme précis ne soit pas complètement élucidé, les protéines gH et gL sont également nécessaires à la réalisation du processus de fusion membranaire et ainsi à l'entrée de la capside virale dans le cytoplasme. Par contre, contrairement à HHV-1 où la glycoprotéine gD est indispensable à la fusion et à la pénétration du virus dans les cellules, gD ne semble pas essentielle lors de la réplication de GaHV-2 (Anderson *et al.*, 1998).

3.2.5.1.2. Transport de la capside au noyau et shutt-off cellulaire.

Après l'entrée du virion au sein du cytoplasme, la migration de la capside jusqu'aux pores nucléaires via le réseau cellulaire de microtubules s'effectue grâce à la protéine de tégument VP1-2 codée par le gène U_L36 (Sodeik *et al.*, 1997).

L'endonucléase VHS (Virion Host Shut-off) codée par le gène U_L41 permet la dégradation des ARNm de la cellule hôte ainsi que ceux du virus. Cette protéine de tégument se localise dans le cytoplasme après l'entrée du virus et permet ainsi une dégradation précoce des ARNm cellulaires par une action similaire à la RNase A. Dans un stade plus avancé de l'infection, elle perdra son activité endonucléasique suite à son association avec la protéine VP16 (Roizman *et al.*, 2007).

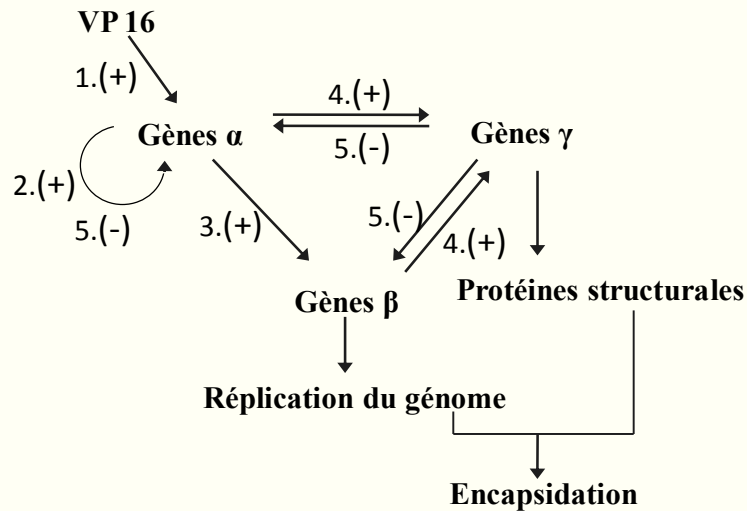


Figure 29 : Cascade des gènes de réplication α, β et γ initiée par la protéine VP16 chez HHV-1. Les chiffres (1, 2, 3, 4 et 5) font référence à la progression du cycle réplcatif.

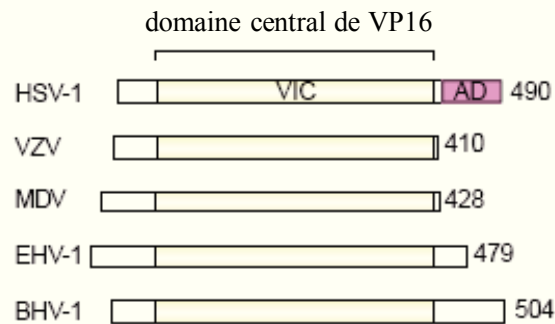


Figure 30 : Comparaison schématique de différentes séquences de VP16 (d'après Wysocka *et al.*, 2003). Ces protéines sont codées par les *Alpha*-herpesvirus HSV-1 (herpes simplex virus type 1), VZV (varicella-zoster virus), MDV (Marek's disease virus), EHV-1 (equine herpesvirus type 1) et BHV-1 (bovine herpesvirus type 1). Le rectangle violet (AD) correspond au domaine C-terminal transactivateur de la protéine de HSV-1. Le domaine central de la protéine (VIC) est conservé entre les différentes séquences. La taille en nombre d'acide aminés est indiquée à droite de chaque séquence.

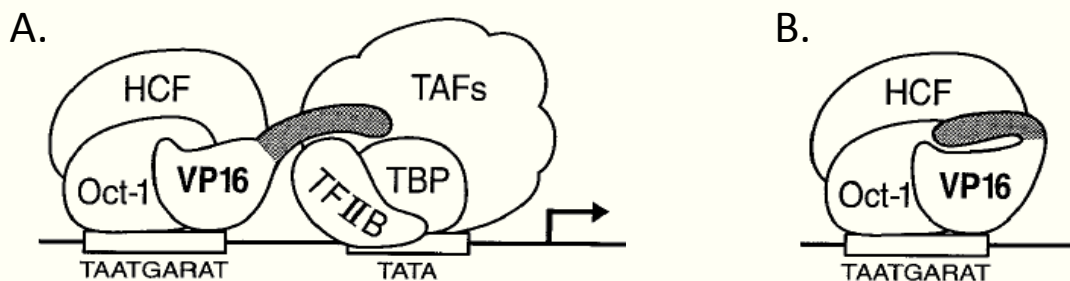


Figure 31 : Représentation schématique du complexe HCF/Oct1/VP16. Complexe formé par la protéine VP16 d'HHV-1 lors de sa fixation sur la séquence consensus TAATGARAT située à proximité du complexe d'initiation de la transcription (A), et à distance de ce complexe (B). La partie grisée de VP16 représente le domaine activateur de la protéine (d'après Hagmann *et al.*, 1997).

3.2.5.1.3. Circularisation de l'ADN viral et Transcription des gènes α .

Dès son entrée dans le noyau de la cellule, l'ADN viral se circularise et il se localise à la périphérie des sites NOD-10 (Nuclear Domain 10) qui correspondent à des sites de défense nucléaire. La chronologie entre l'initiation des gènes viraux et cette localisation nucléaire est encore obscure. En effet, l'initiation de la transcription des gènes très précoces semblerait avoir lieu au niveau des sites NOD-10 (Roizman *et al.*, 2007), alors que la présence des gènes très précoces semble nécessaire à la localisation de l'ADN viral au niveau de ces sites (Roizman *et al.*, 2007). Par ailleurs, il a été montré que la protéine ICP0 (Infected Cellular Protein 0) présente chez HHV-1, mais dont aucun homologue n'a été identifié chez GaHV-2 induit la destruction de ces structures NOD-10.

La protéine tardive VP16, également appelée α -TIF (α -TransInducing Factor) et codée par le gène U_L48 , est un facteur clef de la réplication virale chez HHV-1 (Weinheimer *et al.*, 1992). Initialement présente dans le cytoplasme après le dé-enveloppement du virus, cette protéine va migrer dans le noyau cellulaire indépendamment du virion.

Chez HHV-1, VP16 transactive les gènes α , initiant ainsi la réplication virale (Figure 29). Cette protéine comporte deux domaines fonctionnels, une partie N-terminale impliquée dans l'attachement à l'ADN et une partie C-terminale transactivatrice (Figure 30). La partie N-terminale, relativement conservée entre les protéines VP16 des différents herpesvirus (Figure 30), contient le site de liaison au facteur de transcription Oct-1, impliqué dans la prolifération cellulaire. Le complexe protéique VP16-Oct1 se fixe sur une séquence consensus TAATGARAT (TAAT fixation de Oct1 et GARAT de VP16) généralement retrouvée en amont des gènes α de HHV-1 (Preston *et al.*, 1988). La partie C-terminale, qui n'est pas retrouvée chez VZV et GaHV-2 (Figure 30), est responsable de la transactivation des gènes situés en aval et à proximité du complexe protéique (Figure 31) (Hagmann *et al.*, 1997). Ce domaine permet le recrutement de facteurs clefs du complexe d'initiation de la transcription TBP (TATA Binding protein), TFIIB, TFIID, TFIIA, TFIIH (Flint & Shenk, 1997), et semblerait également favoriser le recrutement de l'ARN Pol II au niveau du complexe de préinitiation de la transcription (Hengartner *et al.*, 1995). Cependant, des études ultérieures ont montré que cette partie COOH terminale transactivatrice n'était pas la région de la protéine la plus impliquée dans la transactivation. En effet, pour toutes les VP16 des herpesvirus, la transactivation se ferait essentiellement via le facteur HCF-1 (Host Cell Factor 1) qui se lie à VP16 par le core de la protéine (Figure 30). Enfin, à la lumière des résultats récents il semblerait que VP16, bien qu'essentielle à l'accomplissement du cycle viral, ne soit pas un facteur indispensable dans cet assemblage, mais plutôt un facteur favorisant l'action du facteur nécessaire HCF-1 (Narayanan *et al.*, 2005). Dans le même esprit, chez GaHV-2, il a été montré que la protéine VP16 n'était pas essentielle pour la réplication virale (Dorange *et al.*, 2002).

Trois gènes très précoces homologues à HHV-1 ont été identifiés chez GaHV-2 et codent pour les protéines ICP4, ICP22 et ICP27. Ces gènes sont les premiers à être exprimés au cours de l'infection et codent pour des protéines impliquées directement ou indirectement dans la transactivation des gènes viraux et cellulaires. Ils vont dans un premier temps s'auto-transactiver, puis permettre la transcription des gènes précoces. Plus tardivement lors du cycle réplcatif, les protéines ICP permettront en collaboration avec les gènes précoces la transcription des gènes tardifs (figure 29). De plus, ces protéines ont un impact sur le déroulement du cycle cellulaire. En effet, elles semblent forcer la cellule à passer en phase S afin d'apporter au virus les éléments nécessaires à sa réplication.

La protéine ICP4, identifiée chez GaHV-2 par Anderson et ses collaborateurs (Anderson *et al.*, 1992), est codée par les gènes *MDV084* et *MDV100*, homologues à *RS1* d'HHV-1, localisés respectivement dans les régions *TR_S* et *IR_S* (figure 27). Cette phosphoprotéine nucléaire de 1415 aa, initierait la transactivation de gènes précoces et tardifs (Anderson *et al.*, 1992). Son pouvoir transactivateur sur son propre promoteur ainsi que sur celui des gènes *pp38*, *pp24* a été démontré par expression ectopique de ICP4 dans la lignée lymphoblastoïde transformée MSB-1 (Endoh *et al.*, 1995; Pratt *et al.*, 1994). Cependant contrairement à son homologue chez HHV-1, ICP4 ne semble pas initier directement la réplication de GaHV-2, car aucune augmentation du nombre de copies de génome de GaHV-2 n'a été retrouvée après transfection du gène *ICP4* (Endoh *et al.*, 1995). De plus, le fait qu'un oligonucléotide « anti-sens » d'*ICP4* entraînerait la perte du potentiel prolifératif de la lignée MSB-1, semblerait suggérer un rôle dans le maintien du phénotype transformé (Xie *et al.*, 1996).

Le rôle précis de la protéine ICP22, codée par *U_S1*, dans la réplication virale de GaHV-2 n'est pas clairement établi. Bien que la réplication *in vitro* d'un virus GaHV-2 délété du gène *U_S1* soit moins importante que celle de la souche sauvage correspondante, et que la coopération ICP22/ICP4 ait un effet positif sur la transcription de la protéine ICP27 (Kato *et al.*, 2002), ICP22 n'est nécessaire ni à l'infection *in vivo*, ni au pouvoir oncogénique de GaHV-2 (Parcells *et al.*, 1995). Chez HHV-1, le gène *U_S1* est également non essentiel, mais il est requis pour une production optimale d'une partie des gènes γ 2. Associée à la protéine kinase codée par *U_L13*, ICP22 modifierait la phosphorylation de l'ARN Pol II favorisant ainsi la transcription des gènes viraux au détriment des gènes cellulaires (Long *et al.*, 1999). ICP22 et *U_L13* permettent également l'activation de *cdc2* et la dégradation de ses partenaires, les cyclines A et B, aboutissant à une expression optimale de certains des gènes γ 2 (Advani *et al.*, 2000a; Advani *et al.*, 2000b).

La phosphoprotéine nucléaire ICP27, codé par le gène *U_L54*, possède une taille de 473 aa et une masse moléculaire de 55 kDa (Ren *et al.*, 1994). De même que son homologue présent chez HHV-1, en fonction du promoteur considéré, cette protéine peut agir en tant qu'activateur ou répresseur. En effet, ICP27 a la capacité d'initier la transcription des gènes codant les protéines *pp38* et *pp24* indépendamment d'ICP4 alors qu'elle réprime l'expression de la thymidine Kinase (TK) (Lupiani *et al.*, 2001). De plus, la protéine ICP27 de GaHV-2 comporte des domaines fonctionnels similaires à ceux de son homologue chez HHV-1 et semble, à l'instar de cette dernière, jouer un rôle dans

l'épissage des gènes viraux et cellulaires au cours de l'infection virale, ce qui permettrait potentiellement de bloquer l'expression de certaines protéines cellulaires ou virales en fonction des étapes du cycle viral (Amor *et al.*, 2011).

3.2.5.1.4. Transcription des gènes β .

Les protéines précoces correspondent à des enzymes impliquées dans la réplication virale et dans le métabolisme des acides nucléiques (Lupiani *et al.*, 2001). Ces enzymes sont très conservées et ont été identifiées par analogie de séquence avec HHV-1. Sept d'entre elles sont nécessaires à la réplication virale : l'ADN polymérase (U_L30) et son facteur de processivité (U_L42), une protéine se liant à l'origine de réplication virale (U_L9 , ori binding protein), une protéine liant l'ADN simple brin (U_L29 ou ICP8) ainsi que le complexe hélicase-primase formé de trois protéines (U_L5 , U_L8 et U_L52). Bien que non-essentiels pour le déroulement de base de la réplication du génome, la thymidine kinase (U_L23), le complexe ribonucléotide réductase (U_L39 , U_L40), l'uracil DNA glycosylase (U_L2), la deoxyribonucléase (U_L12) et la dUTPase (U_L50) restent des enzymes importantes pour la virulence, l'infectiosité et la réactivation du virus.

3.2.5.1.5. Réplication ADN viral par les gènes β .

Une fois le génome viral circularisé, la protéine U_L9 se lie à l'origine de réplication permettant le déroulement de l'ADN puis le recrutement d'ICP8 ainsi que de différents facteurs venant se fixer au niveau de la fourche de réplication. Par la suite, l'assemblage des complexes hélicase-primase et polymérases virales avec ce premier complexe est à l'origine de l'initiation du cycle de réplication théta. Ensuite, le passage de la forme répllicative théta à la réplication par cercle roulant s'effectue suite au départ de la protéine U_L9 . Ce mode de réplication permet de générer des génomes viraux linéaires sous forme de concatémères qui seront clivés en structures monomériques lors de l'assemblage des virions. Il faut noter que des fourches de réplication secondaires sont très souvent décrites, donnant à l'ADN en cours de réplication un aspect branché (Severini *et al.*, 1996).

3.2.5.1.6. Expression des gènes γ .

Les protéines tardives correspondent classiquement aux protéines de structure du virus et sont rassemblées en trois groupes : les glycoprotéines d'enveloppe, les protéines de capsid et les protéines de tégument. Ces protéines sont activées en fin de cycle viral par les gènes très précoces notamment par *ICP4* ainsi que par les gènes précoces.

3.2.5.1.6.1. Les glycoprotéines

GaHV-2 code pour dix glycoprotéines différentes homologues à HHV-1, dont sept sont codées par des gènes retrouvés dans U_L: gL (U_L1), gM (U_L10), gH (U_L22), gB (U_L27), gC (U_L44), gN (U_L49.5), gK (U_L53), et trois dans U_S: gD (U_S6), gI (U_S7), gE (U_S8) (Brunovskis & Velicer, 1995; Lee *et al.*, 2000; Tulman *et al.*, 2000). De manière générale, les glycoprotéines des herpesvirus sont impliquées dans les processus initiaux d'attachement, de fusion membranaire, de pénétration du virion, de trafic intracellulaire des composants du virion, de la sortie et la diffusion de cellule à cellule du virus (Rajcani & Vojvodova, 1998). Le rôle prépondérant de la plupart des glycoprotéines dans les interactions virus/cellules a été mis en évidence par le développement de stratégies d'inactivation de gènes viraux au sein des génomes de GaHV-2. Ainsi, pour les glycoprotéines de GaHV-2, la délétion de gB ou de l'une ou l'autre des protéines formant les hétérodimères gE/gI et gM/gN entraîne une incapacité de multiplication et de diffusion du virus *in vitro* (Schumacher *et al.*, 2000; Schumacher *et al.*, 2001; Tischer *et al.*, 2002). En plus de son rôle dans la dissémination de cellule à cellule, le complexe gH/gL joue également un rôle important dans les étapes de pénétration du virus (Wu *et al.*, 2001). A l'instar des autres herpesvirus, gH doit interagir avec gL pour permettre leur translocation à la surface des cellules (Wu *et al.*, 2001; Wu *et al.*, 1999). La glycoprotéine gC, qui est retrouvée très exprimée au niveau de la surface cellulaire et du cytoplasme des cellules infectées, semble fortement impliquée dans le pouvoir infectieux du virus, car indispensable pour la transmission horizontale de GaHV-2 (Jarosinski *et al.*, 2007; Jarosinski & Osterrieder, 2010). Il faut noter que les glycoprotéines gB et gC, correspondent à l'antigène B associé aux cellules et à l'antigène soluble A, initialement mis en évidence par Churchill et Biggs par un test d'immunodiffusion en gélose (Churchill & Biggs, 1967). Historiquement ces antigènes ont été largement utilisés dans des tests de diagnostic de la maladie de Marek. Contrairement à HHV-1, la glycoprotéine gD de GaHV-2 ne semble pas essentielle à la fusion et à la pénétration du virus dans les cellules. En effet, un virus invalidé pour cette protéine conserve son pouvoir réplcatif *in vitro* et oncogénique *in vivo* ainsi que sa capacité à diffuser au sein des élevages (Anderson *et al.*, 1998). Enfin, la glycoprotéine gK a été caractérisée chez GaHV-2, cependant son rôle n'a pas été démontré comme chez HHV-1 chez lequel la protéine homologue interagit directement avec gB lors des phases de fusion membranaire et de sortie du virus (Chouljenko *et al.*, 2010).

3.2.5.1.6.2. Les protéines du tégment

Chez HHV-1, quatre des quinze protéines de tégment (VP11/12, VP13/14, VP16 et VP22) sont des composants majeurs des virions et sont susceptibles d'intervenir dans la morphogénèse et dans la régulation de l'expression des gènes viraux. Les autres protéines composant le tégment possèdent des fonctions très variées (activité kinase, blocage de la synthèse de protéines cellulaires, interaction avec les ribosomes et encapsidation de l'ADN).

Les gènes U_L46, U_L47, U_L48 et U_L49 codant respectivement pour les protéines VP11/12, VP13/14, VP16 et VP22, sont également présents chez GaHV-2 (Yanagida *et al.*, 1993). Cependant, il a été montré que seul VP22 semble indispensable à la réplication de GaHV-2 *in vitro* (Dorange *et al.*, 2002). De plus elle est essentielle pour la diffusion des particules virales de cellule à cellule (Blondeau *et al.*, 2008).

Par ailleurs, la VP22 d'HHV-1 présente également de nombreuses fonctions. Elle joue un rôle majeur dans la réorganisation des microtubules (Elliott & O'Hare, 1998). Elle est également capable d'interagir avec gD pouvant permettre une liaison de gD à la capside du virion (Chi *et al.*, 2005) et avec ICP0 afin de limiter le pouvoir transactivateur de ce dernier sur les gènes de la thymidine kinase et de gC (Yu *et al.*, 2010). Enfin, VP22 ainsi que VP11/12 et VP13/14 semblent également capables de moduler les fonctions transactivatrices de VP16 (Elliott *et al.*, 1995; Zhang *et al.*, 1991).

3.2.5.1.6.3. Les protéines de capside

Les capsides virales matures à symétrie icosaédrique sont composées de 162 capsomères et mettent en jeu 7 protéines : la protéine majeure de capside VP5 (U_L19), la protéine d'échafaudage pré-VP22a (U_L26.5), les protéines VP19c (U_L38), VP23 (U_L18), VP26 (U_L35), et les protéines VP21-VP24 issues de l'autoclivage de la protéase codée par le gène U_L26. Comme chez HHV-1 (Tatman *et al.*, 1994), il a été montré qu'une coexpression ectopique de VP5, preVP22a, VP23, VP19c de GaHV-2 dans un système baculovirus était nécessaire et suffisante pour leur auto-assemblage sous forme de pseudocapsides virales présentant une morphologie identique à des capsides immatures natives (Kut & Rasschaert, 2004).

3.2.5.1.6.4. La lipase virale

La lipase virale possède de fortes homologies de séquence en acides aminés avec les motifs α/β hydrolase de lipases pancréatiques cellulaires (Kamil *et al.*, 2005). Ce gène spécifique du genre *Mardivirus* est issu d'un gène tardif qui est exprimé spécifiquement durant la phase lytique. Il est composé de deux exons, un premier codant pour le peptide signal alors que le second code pour les domaines à activité lipase (Kamil *et al.*, 2005). Bien que non essentielle, vLIP (Viral Lipase) permet d'augmenter la réplication virale ainsi que le pouvoir pathogène de GaHV-2 (Kamil *et al.*, 2005). Cependant cette glycoprotéine sécrétée de 756 aa ne semble pas posséder d'activité hydrolase comme ces homologues cellulaires et son rôle précis reste inconnu.

3.2.5.1.7. Assemblage de la capside.

Après leur synthèse, la protéine majeure de capside VP5 (U_L19), les protéines de capside VP26 (U_L35), VP19c et VP23, ainsi que les protéines d'échafaudages preVP22a (U_L26.5) et VP24-

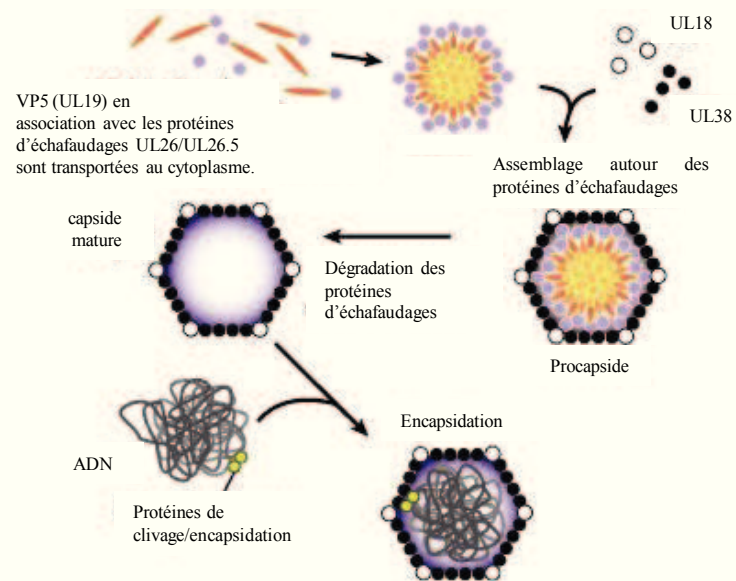


Figure 32 : Assemblage de la capside et encapsidation de l'ADN viral chez HHV-1.

VP21 (U_L26) sont importées dans le noyau, siège de l'assemblage des virions. Autour de la structure d'échafaudage formée par les protéines VP22A et VP24/VP21 les pentamères de VP5 et les hexamères de VP5-VP26 viennent s'assembler avec les triplex de connexion constitués de l'assemblage des protéines VP19c-VP23, pour former une structure relâchée stable : la procapside (Figure 32). Ensuite, la séparation par autoclivage de la protéase VP24 et de la protéine VP21 conduit à la perte de l'échafaudage (Figure 32). A ce moment, la capside virale mature est prête à intégrer l'ADN viral.

L'assemblage de la capside de GaHV-2 est identique à celui des autres herpesvirus, hormis l'autoclivage de la protéase au niveau des sites R et M permettant de générer VP24 (protéase) d'une part et VP21 d'autre part, qui s'effectue de manière séquentielle et non simultanée. En effet, contrairement à tous les autres herpesvirus, le site R est le premier à être ciblé par l'activité autocatalytique de la protéase (Laurent *et al.*, 2007) et la VP24 séparée de VP21 mature ensuite cette dernière par élimination des quelques aa de l'extrémité COOH terminale par coupure au niveau du site M. Il faut noter également que le site M, bien que fonctionnel, possède une séquence en acides aminés non consensuelle (Laurent *et al.*, 2007).

3.2.5.1.8. Encapsidation.

Le clivage et l'encapsidation d'une unité génomique au sein de la capside virale est un processus clef dans le cycle des herpesvirus. Le mécanisme ainsi que le rôle des protéines impliquées dans le clivage et l'encapsidation sont peu décrits chez GaHV-2. Cependant, du fait que ce mécanisme soit très conservé au sein des herpesvirus, et que les protéines de GaHV-2 présentent de fortes homologues avec les protéines d'HHV-1, il est supposé un rôle commun de ces protéines entre les deux virus.

Chez HHV-1, l'encapsidation de l'ADN viral nécessite sept protéines : U_L6, U_L15, U_L17, U_L25, U_L28, U_L32 et U_L33. Les protéines U_L15, U_L28 et U_L33 s'associent pour former le complexe terminase qui permet la fragmentation du concatémère d'ADN, en unités génomiques simples. Il induit des coupures de l'ADN double brin au niveau de sites spécifiques du concatémère appelés séquences α , de telle façon qu'un génome viral complet puisse être introduit dans chaque capside. Les séquences α sont constituées de deux régions distinctes indispensables au mécanisme de clivage et d'encapsidation : *pac1* et *pac2* situées de part et d'autre du site de clivage. Chez GaHV-2, ces séquences se situeraient en amont du transcrit LAT (Latency Associated Transcript), au niveau des régions répétées IR_S et TR_S à proximité de la jonction avec l'IR_L ou TR_L.

Au sein du complexe terminase, la protéine U_L15 possède une activité ATPase nécessaire à l'encapsidation. En effet, le mécanisme d'encapsidation est particulièrement consommateur d'énergie, puisqu'il a été estimé qu'une molécule d'ATP était hydrolysée pour deux pb intégrées. De plus, la protéine UL15 est un facteur clef dans la localisation d'UL28 et UL33 au niveau des centres de

réplication où a lieu l'encapsidation (Higgs *et al.*, 2008). L'interaction entre le complexe terminase et l'ADN s'effectue entre la protéine U_L28 et une séquence *pac1*. Par ailleurs, les protéines U_L15 et U_L28 interagissent également avec la protéine U_L6, qui forme le site d'entrée de l'ADN dans la capside (White *et al.*, 2003). Plus précisément, douze monomères de la protéine UL6 en interagissant spécifiquement avec l'un des douze sommets de la capside forment un anneau définissant un canal permettant l'entrée de l'ADN dans la capside (Newcomb *et al.*, 2001).

Parmi les autres protéines impliquées dans le mécanisme de clivage/encapsidation, la protéine U_L17, également essentielle chez GaHV-2 (Chhab *et al.*, 2009), semble être nécessaire à la localisation des capsides au niveau des sites intranucléaires de réplication (Salmon *et al.*, 1998; Taus *et al.*, 1998). La protéine UL32 qui ne se lie pas à la capside, serait également impliquée dans le transfert des capsides préformées vers les centres de réplication (Lamberti & Weller, 1998). Enfin, la protéine U_L25 est requise pour une encapsidation entière du génome viral. Elle ne semble pas impliquée dans le clivage de l'ADN viral (McNab *et al.*, 1998), mais elle empêcherait l'échappement de l'ADN encapsidé en le stabilisant à l'intérieur de la capside virale (Ogasawara *et al.*, 2001).

A ces différentes protéines, nous pouvons également ajouter la nucléase alcaline codée par U_L12, qui n'est pas essentielle mais facilite le clivage et l'encapsidation de l'ADN viral. Un virus mutant pour la nucléase alcaline possède un titre viral 100 à 1000 fois plus faible que son homologue sauvage (Shao *et al.*, 1993). Cette baisse de titre serait due à la somme de diminutions ponctuelles de rendement lors de la synthèse de l'ADN viral, de son implication dans l'encapsidation, de la sortie du virus du noyau et de la capacité des nouvelles particules à initier de nouveaux cycles réplcatif (Porter & Stow, 2004).

3.2.5.1.9. Sortie du virus.

Après l'encapsidation de l'ADN viral, les nucléocapsides formées sont capables de bourgeonner à travers la membrane interne du noyau. Ce processus implique des interactions entre la nucléocapside, les protéines du tégument et certaines glycoprotéines virales, dont essentiellement gK. Le virus à ce moment n'aurait acquis qu'une partie du tégument. En effet, ce dernier apparaît moins dense en microscopie électronique, que celui des virions matures après bourgeonnement hors de la cellule, laissant ainsi supposer des modifications ultérieures (Roizman *et al.*, 2007). La traversée de la membrane nucléaire par les particules intermédiaires s'effectue grâce au complexe nucléaire de sortie constitué des protéines U_L31 et U_L34 en collaboration avec la kinase virale U_S3 et plusieurs kinases cellulaires, via une interaction avec les lamines. Après fusion de la membrane intermédiaire avec le feuillet externe de la membrane nucléaire, la nucléocapside entourée du tégument rejoint le réseau trans-golgien soit via une vésicule, soit via le réticulum endoplasmique. Il faut noter que l'enveloppe définitive ainsi que le reste des protéines de tégument sont acquis par bourgeonnement durant ces étapes.

Par ailleurs, en plus de se propager de cellule à cellule via le compartiment extracellulaire, certains herpesvirus comme HHV-1 et GaHV-2, peuvent passer directement de cellule à cellule, via l'implication des glycoprotéines gE et gI.

3.2.5.2. Transcriptome des régions TR_L/IR_L et IR_S/TR_S de GaHV-2.

Les gènes spécifiques de GaHV-2 sont principalement retrouvés au niveau des régions répétées inversées IR_L/TR_L et IR_S/TR_S et donc présents en deux copies dans le génome. Bien que certains soient communs au genre *Mardivirus*, la plupart sont spécifiques de chaque sérotype.

3.2.5.2.1. Les microARN de GaHV-2

Après la découverte de miARN chez la plupart des herpesvirus (Pfeffer *et al.*, 2005; Pfeffer *et al.*, 2004), différentes études ont montré la présence de 13 pré-miARN (pré-miR) permettant l'expression de 25 miARN matures chez GaHV-2 (Burnside *et al.*, 2006; Burnside *et al.*, 2008; Morgan *et al.*, 2008; Yao *et al.*, 2008). Ces miARN ont été regroupés en deux clusters. Un premier, localisé dans l'IR_L/TR_L, encadre l'oncogène *meq* (Figure 33). Ce cluster est constitué des miR-M9, M5, M12, M3, M2 et M4, situés en amont de *meq* et des miR-M11, M31 et M1 localisé en aval (Figure 33). Le second cluster, constitué des miR-M8, M6, M7 et M10, est localisé dans la partie 5' du transcrit LAT au niveau de l'IR_S/TR_S (Figure 34). Plusieurs banques de miARN ont été réalisées à partir de cultures de CEF infectées par GaHV-2 (Burnside *et al.*, 2006), de lymphomes de rate induits par GaHV-2 (Morgan *et al.*, 2008), de la lignée lymphocytaire CD4+ MSB-1 transformée par la souche vBC-1 (Morgan *et al.*, 2008; Yao *et al.*, 2008), de la lignée lymphocytaire 54-O transformée par la souche vvRB-1B, ainsi que de PBL (Peripheral Blood Leucocytes) à J5, J21, J31 et de lymphocytes T CD4+ triés à J21, récupérés lors d'une infection *in vivo* par la souche vvRB-1B (données du laboratoire non publiées. B.Muylkens). L'analyse comparative de ces banques fait ressortir des différences d'expression de ces miARN conduisant à suspecter une régulation spécifique en fonction des types de cellules infectées, du cycle viral et du pouvoir pathogène du virus.

Ces différences sont déjà observables dans l'expression des différents miARN rapportée à celle des miARN viraux totaux. Quelques miARN, tels que mdv1-miR-M4-5P, mdv1-miR-M3-5P ou mdv1-miR-M5-3P, sont fortement exprimés dans l'ensemble des banques (Tableau 4). Certains semblent plus spécifiques de la phase lytique, tels que mdv1-miR-M6-3P, mdv1-miR-M8-3P, mdv1-miR-M2-5P ou mdv1-miR-M1-5P, et d'autres plus spécifiques de la phase de latence et de transformation tels que, mdv1-miR-M3-5P, mdv1-miR-M6-5P, mdv1-miR-M7-5P ou mdv1-miR-M8-5P (Tableau 4). Il apparaît également une forte expression du cluster 1 par rapport au cluster 2 lors de la phase lytique, et une proportion égale entre les deux clusters pendant la phase de latence et la

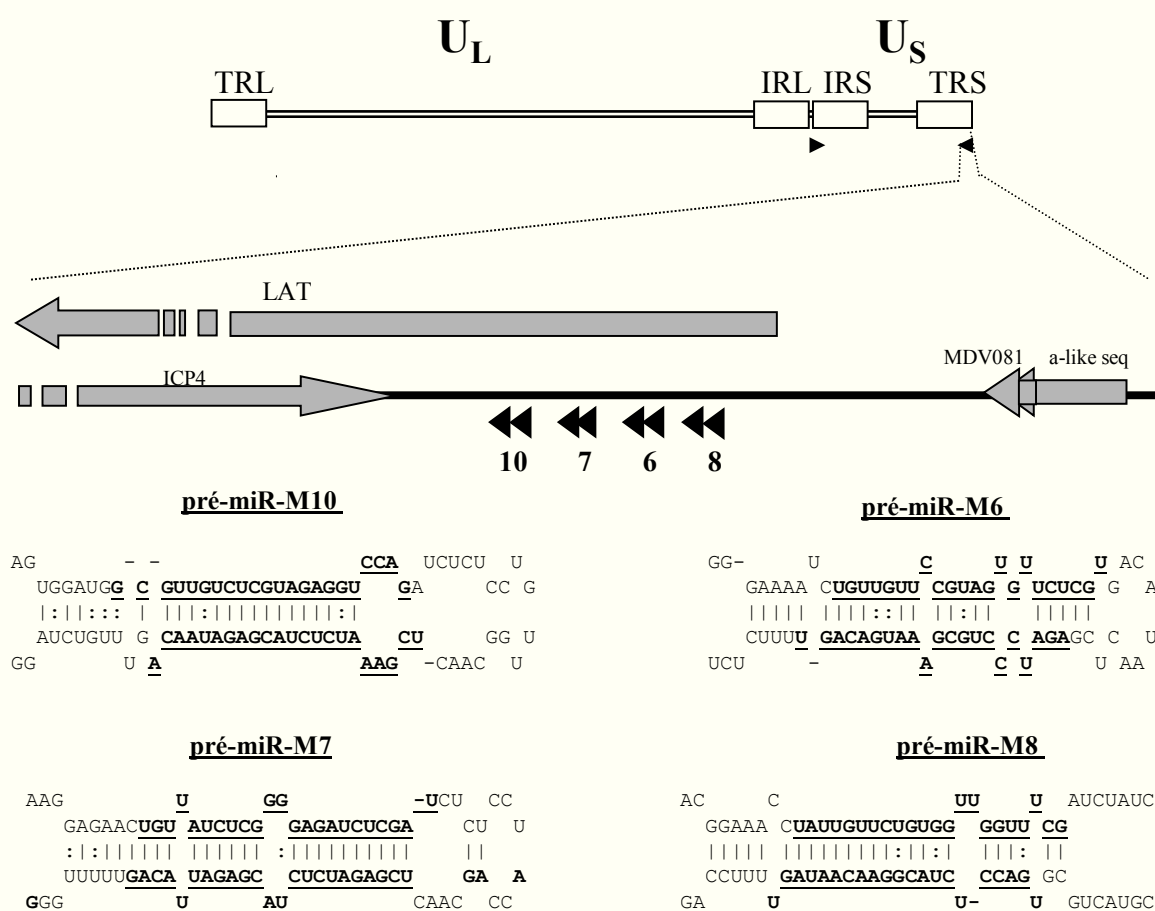


Figure 34 : Localisation et représentation de la structure tige-boucle des 4 pré-microARN issus du cluster 2 de microARN, localisés dans les régions IR_S et TR_S de GaHV-2.

Les microARN matures 5P et 3P sont en caractère gras et soulignés au niveau du brin supérieur (5P) et du brin inférieur (3P) de la structure en tige boucle. Les liaisons fortes entre A-U et G-C sont représentées par un trait tandis que les liaisons plus faibles entre G et U sont représentées par deux points.

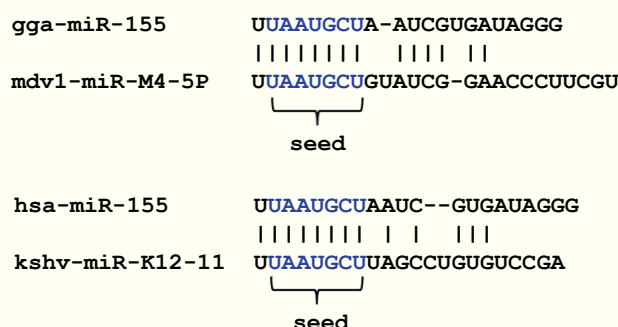


Figure 35 : Alignement des séquences des miR-155 cellulaires humain (hsa) et de *gallus gallus* (gga) avec leurs orthologues viraux HHV-8 (KSHV) et GaHV-2.

	CEF infecté 24h (Burnside et al., 2006)	CEF infecté 24h (Burnside et al., 2008)	tumeur de rate (Morgan et al., 2008)	MSB-1 [*] (Yao et al., 2008)	MSB-1 [*] (Morgan et al., 2008)	MSB-1 [§]	54-O [§]	PBL J 21 [§]	lymphocytes T CD4+ J 21 [§]	PBL J 31 [§]
total miR GaHV-2	141	1036	13209	2562	245475	782	551	160	86	237
% miARN GaHV-2 [¶]	1,0	1,0	0,4	50,3	10,0	41,5	38,7	10,0	4,3	14,9
% microARN cluster 1	89,2	79,9	96,1	55,0	75,3	48,3	47,1	20,7	67,4	19,8
% microARN cluster 2	10,8	20,1	3,8	45,0	24,9	51,8	53,0	79,3	32,6	79,6
mdv1-miR-M9-5P		0,5	0,8	1,8	0,1	1,4	0,5	1,3	2,3	
mdv1-miR-M9-3P			1,0	0,0	0,1	0,4				0,4
mdv1-miR-M5-5P		1,2	< 0,1		0,0	0,4	0,5			
mdv1-miR-M5-3P	5,6 [¶]	6,0	6,3	6,9	33,0	2,2	1,6		3,5	0,4
mdv1-miR-M12-5P			0,2		0,0	0,1				
mdv1-miR-M12-3P		0,3	2,3	1,4	2,3	11,3	9,4	6,0	16,3	5,5
mdv1-miR-M13-5P	2,9	1,3	3,9	15,2	2,8	14,1	17,1	5	11,6	5,5
mdv1-miR-M13-3P			< 0,1		0,0		0,4			
mdv1-miR-M12-5P		18,4	1,6	0,6	1,1	0,9	1,1	0,6	2,3	
mdv1-miR-M12-3P	3,6	1,5	6,9	0,4	21,5	0,6				
mdv1-miR-M14-5P	28,8	19,9	72,0	13,2	7,6	14,6	14,9	6,3	27,9	7,6
mdv1-miR-M14-3P	2,2	0,6	0,1	2,0	0,4	1,9	1,1	0,6	1,2	0,4
mdv1-miR-M11-5P				0,1						
mdv1-miR-M11-3P		1,0								
mdv1-miR-M31-3P			0,0		0,1					
mdv1-miR-M1-5P	23,0	29,3	1,0	13,2	3,6	0,4	0,5		2,3	
mdv1-miR-M1-3P			0,0	0,2						
mdv1-miR-M8-5P		1,1	3,5	0,7	23,5	4,7	3,1	2,5	9,3	5
mdv1-miR-M8-3P	2,2	4,8	0,1	1,4	0,9	0,5	1,3	1,3	1,2	
mdv1-miR-M16-5P		3,8	< 0,1	10,9	0,1	6,8	8,2	1,9	3,5	2,5
mdv1-miR-M16-3P	7,2	8,4	0,1	0,5	0,1		0,4	0,6		
mdv1-miR-M7-5P				31,2	0,3	39,6	39,6	73	18,6	71,7
mdv1-miR-M7-3P	1,4	1,4			0,0	0,1	0,2			0,4
mdv1-miR-M10-5P		0,4	0,1		0,0		0,2			
mdv1-miR-M10-3P		0,2		0,2	0,0	0,1				

Tableau 4 : Expression des 25 microARN matures retrouvés dans les différentes banques issues de la littérature ou des banques obtenues au sein du laboratoire.

* Les cellules MSB-1 contiennent également environ 10 % de microARN de GaHV-3,

§ banques issues du laboratoire,

¶ les cellules 54-O sont une lignée lymphocytaire CD4+ transformée par la souche GaHV-2 RB-1B issue d'une tumeur ovarienne,

‡ les banques PBL (Peripheral Blood Leucocytes) J21, J31 et lymphocytes T CD4+ sont issues d'une infection *in vivo* par une souche vvRB-1B,

¶ le pourcentage de miARN de GaHV-2 (% miR GaHV-2) correspond au nombre de microARN de GaHV-2 rapporté à l'ensemble des petits ARN séquencés (total miR GaHV-2) au sein de chaque banque. Le pourcentage indiqué pour les différents miARN est relatif à l'ensemble des séquences issues de GaHV-2,

¶ le pourcentage d'un microARN mature présent à un taux supérieur à 5% est indiqué en rouge,

† SNP identifié entre la souche GaHV-2 RB-1B et la souche GaHV-2 BC-1.

transformation (Tableau 4). Il faut noter que selon les auteurs, les lignées ne semblent pas être homogènes quant à leur expression des miARN. En effet, contrairement aux résultats obtenus par notre équipe et l'équipe de Yao, il apparaît que les cellules MSB-1 de Morgan et ses collaborateurs expriment mdv1-miR-M5-3P et mdv1-miR-M8-5P à un taux très élevé tandis que mdv1-miR-M7-5P n'est que très peu représenté dans ces cellules.

Parmi les 25 miARN clonés, seules des cibles ont été attribuées à mdv1-miR-M3-5P et mdv1-miR-M4-5P. Le mdv1-miR-M3 semble avoir des propriétés anti-apoptotiques puisqu'une étude récente a montré que ce miARN diminue le taux de mortalité des cellules traitées par cisplatine, molécule utilisée en traitement anticancéreux visant à induire l'apoptose (Xu *et al.*, 2011). Cette propriété anti-apoptotique de mdv1-miR-M3 est corrélée à la régulation directe de l'expression de la protéine SMAD 2 qui fait partie de la voie de signalisation du TGF- β . Le mdv1-miR-M4-5P, fortement exprimé dans l'ensemble des banques, possède une séquence « seed » (nucléotides 2-8), identique à celle d'un miARN cellulaire, le miR-155 (Figure 35). Ce dernier, que l'on peut qualifier d'oncomiR, possède un rôle clef dans de nombreux processus physiologique et pathologique de la biologie des lymphocytes. Ainsi, de par leur homologie de séquence, mdv1-miR-M4 et *Gallus gallus* gga-miR-155 ciblent et dérèglent un même ensemble de gènes cellulaires (PU.1, CEBP β , HIVP2, BCL2L13, PDCD6) (Zhao *et al.*, 2009b). De plus, il est intéressant de noter que le gammaherpesvirus de Kaposi HHV-8 code également pour un orthologue de miR-155, le kshv-mir-K12-11-3P (Figure 35) (Gottwein *et al.*, 2007; Skalsky *et al.*, 2007). Comme son homologue humain hsa-miR-155, ce miARN viral cible et dérègle les gènes cellulaires *BACH1* (Skalsky *et al.*, 2007), *c-MAF* (Hansen *et al.*, 2010; Rodriguez *et al.*, 2007), *FOS*, *MAP3K10*, *AGTRAP*, *BIRC4BP*, *IKBKE*, *TWF1*, *PIK3CA*, *RFK*, *RPS6KA3*, *SAMHD* et *SLA* (Gottwein *et al.*, 2007). De plus, kshv-mir-K12-11 en association avec d'autres miARN de HHV-8 et notamment kshv-mir-K12-6, diminue l'expression du suppresseur de tumeur *THBS1* (Samols *et al.*, 2007) et du facteur de transcription *c-MAF* (Hansen *et al.*, 2010). Cependant, à l'instar de mdv1-miR-M4, aucun rôle dans la régulation directe de gènes viraux n'a été mis en évidence.

Des miARN ont également été retrouvés chez les autres *Mardivirus*. En effet, 17 pré-miR ont été identifiés chez GaHV-3 (HPRS-24) (Yao *et al.*, 2007), et 18 pré-miR chez MeHV-1 (HVT) (Waidner *et al.*, 2009; Yao *et al.*, 2009b). Cependant, bien qu'ils soient également localisés dans les régions répétées du génome, aucune homologie n'a été mise en évidence entre les miARN des différents sérotypes. Il faut noter que pour la première fois chez un virus, le virus HVT, il a été montré une évolution des miARN à partir d'un évènement de duplication (Yao *et al.*, 2009b).

Enfin, comme pour tous les virus, la comparaison des banques de miARN cellulaires de tissus sains ou infectés, fait apparaître que l'infection de cellules ou de poulets par GaHV-2 induit une dérégulation de nombreux miARN cellulaires. (Burnside *et al.*, 2008; Yao *et al.*, 2009a). De plus, les miARN dont les taux varient lors de la phase lytique sont différents de ceux variant lors des phases de latence/oncogenèse. Ainsi, lors d'une infection lytique de CEF par GaHV-2, gga-miR-10b, miR-29b,

miR-30d, miR-133a et miR-196 sont surexprimés tandis que l'expression de gga-miR-1a, miR-130a, let-7a, let-7b, let-7f est réprimée (Burnside *et al.*, 2008), alors que dans les lignées MSB-1, les miR gga-miR-7b, miR-18a, miR-29a, miR-34a, miR-133a, miR-146b, miR-454 et let-7d sont retrouvés surexprimés et les miR gga-miR-26a, miR-126, miR-150, miR-155, miR-223 et miR-451 sous-exprimés (Yao *et al.*, 2009a).

3.2.5.2.2. Transcriptome, protéome de IRS/TRS

Seulement deux ARN sont transcrits à partir de cette région IRS/TRS, le messager d'ICP4 sur un brin et le transcrit LAT (Latency Associated Transcript) dans l'autre brin. L'analyse comparative du profil d'expression des transcrits LAT et d'*ICP4* indique une transcription importante du gène *ICP4* lors de l'infection cytotytique en CEF alors que les ARN du *LAT* seraient majoritairement détectés dans les lignées cellulaires transformées par GaHV-2 et dans les cellules tumorales. Le rôle de la protéine ICP4 ayant été détaillé précédemment, je ne présenterai ici que la famille des transcrits LAT.

La présence d'ARN non codants tels que les LAT chez HHV-1, les EBER chez HHV-4 ou encore l'ARN $\beta 2.7$ chez HHV-5 (Sullivan, 2008) a été rapportée pour plusieurs herpesvirus. Chez GaHV-2, une famille de transcrits associés à la latence (LAT) a été identifiée sur le brin anti-sens du gène codant la protéine ICP4 (Cantello *et al.*, 1994). Cette famille de transcrits a été retrouvée *in vivo* dans des lymphomes induits par GaHV-2, dans les différentes lignées lymphoblastoïdes établies à partir de lymphomes induits par GaHV-2 et *in vitro* lors d'infections cytotytiques de CEF par GaHV-2 (Cantello *et al.*, 1994; Cantello *et al.*, 1997; Li *et al.*, 1998; Li *et al.*, 1994). Elle comprend de nombreux transcrits alternatifs polyadénylés ou non, dont un grand ARN de 11 kb non épissé, polyadénylé recouvrant la presque totalité de l'IRS ou TRS, ainsi que trois grandes familles de transcrits alternatifs (Figure 36), parmi lesquels ont été décrits des petits ARN nucléaires épissés non polyadénylés de 0,75 et 0,9 kb appelés MSR (MDV Small RNA), ainsi que de nombreux ARN hautement épissés et polyadénylés (Cantello *et al.*, 1994; Cantello *et al.*, 1997; Li *et al.*, 1998; Li *et al.*, 1994) ; communication personnelle. S. Laurent).

A ce jour, aucune protéine n'a été associée à ces différents transcrits LAT composés de 15 exons alternatifs dont 9 possèdent des bornes variables (communication personnelle. S. Laurent) et seul le cluster 8-10 de miARN a été identifié dans la partie 5' du *LAT* (Burnside *et al.*, 2006). Plus précisément, ce cluster de miARN est présent dans le premier intron des trois groupes de transcrits alternatifs LAT qui s'initient dans ou juste en aval de la dernière des répétitions de 60 pb (Figure 36 ; communication personnelle. S. Laurent) (Stik *et al.*, 2010). Par ailleurs, des transcrits analogues ont été retrouvés chez HHV-1, sur le brin complémentaire du gène *ICP0*, recouvrant la totalité de ce dernier. De plus, des miARN et deux petits ARN (sRNA : small RNA) permettant, entre autres, l'inhibition de l'expression du gène *ICP0* et ou *ICP4*, ont également été identifiés dans cette région

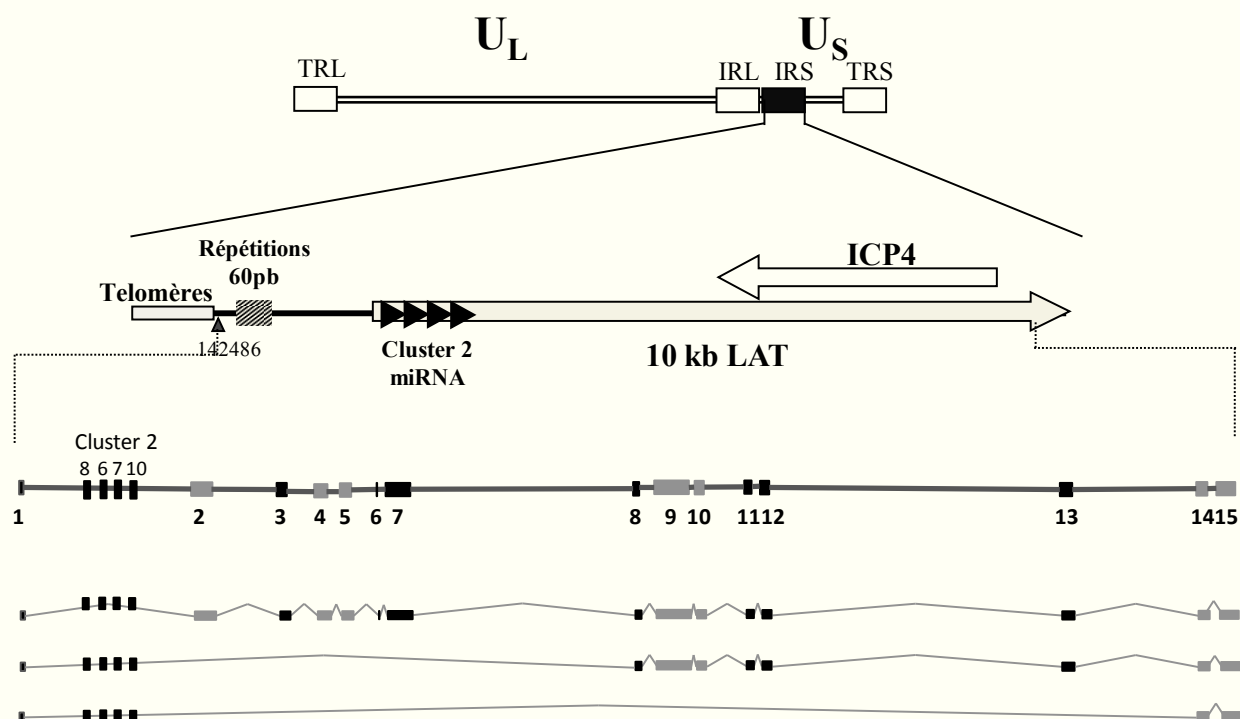


Figure 36 : Représentation schématique des trois familles de transcrits alternatifs issus du gène LAT. Les familles de transcrits sont définies par leur épissage : exon1-exon2, exon1-exon8 et exon1-exon14. Les microARN du cluster 2 sont représentés en rouge, les exons du gène *LAT* sont représentés en gris s'ils possèdent des bornes d'épissage variables et en noir s'ils possèdent des bornes d'épissage fixes. Les répétitions de 60 pb, présentes en nombre variable en fonction de la souche sont représentées en gris hachuré.

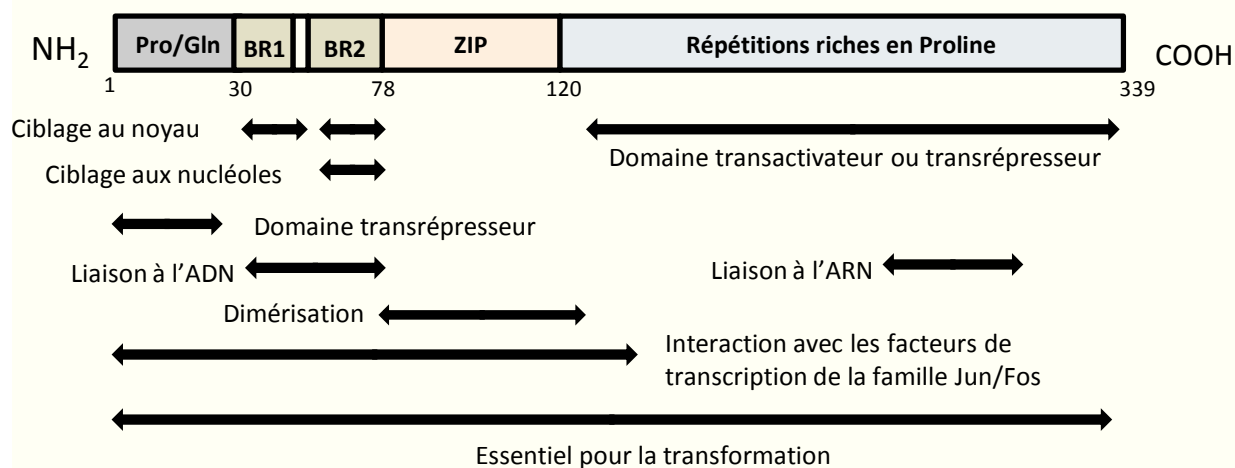


Figure 37 : Structure et localisation des différents domaines de la protéine Meq du virus GaHV-2. La fonctionnalité des domaines nécessaires pour les interactions protéines/protéines et protéines/acides nucléiques, ainsi que les domaines impliqués dans la localisation nucléaire et nucléolaire sont indiqués par des flèches noires.

(Shen *et al.*, 2009; Umbach *et al.*, 2008). Ces résultats laissent à penser que les LAT de GaHV-2 pourraient réguler négativement la transcription du gène *ICP4* par une régulation par ARN anti-sens. Mais surtout, au travers des miARN du cluster 8-10, l'expression du transcrit LAT pourrait réguler l'expression de différents gènes cellulaires et viraux permettant d'une part l'installation ainsi que le maintien de la latence et conduisant d'autre part au développement des tumeurs.

3.2.5.2.3. Transcriptome, protéome de IR_L/TR_L

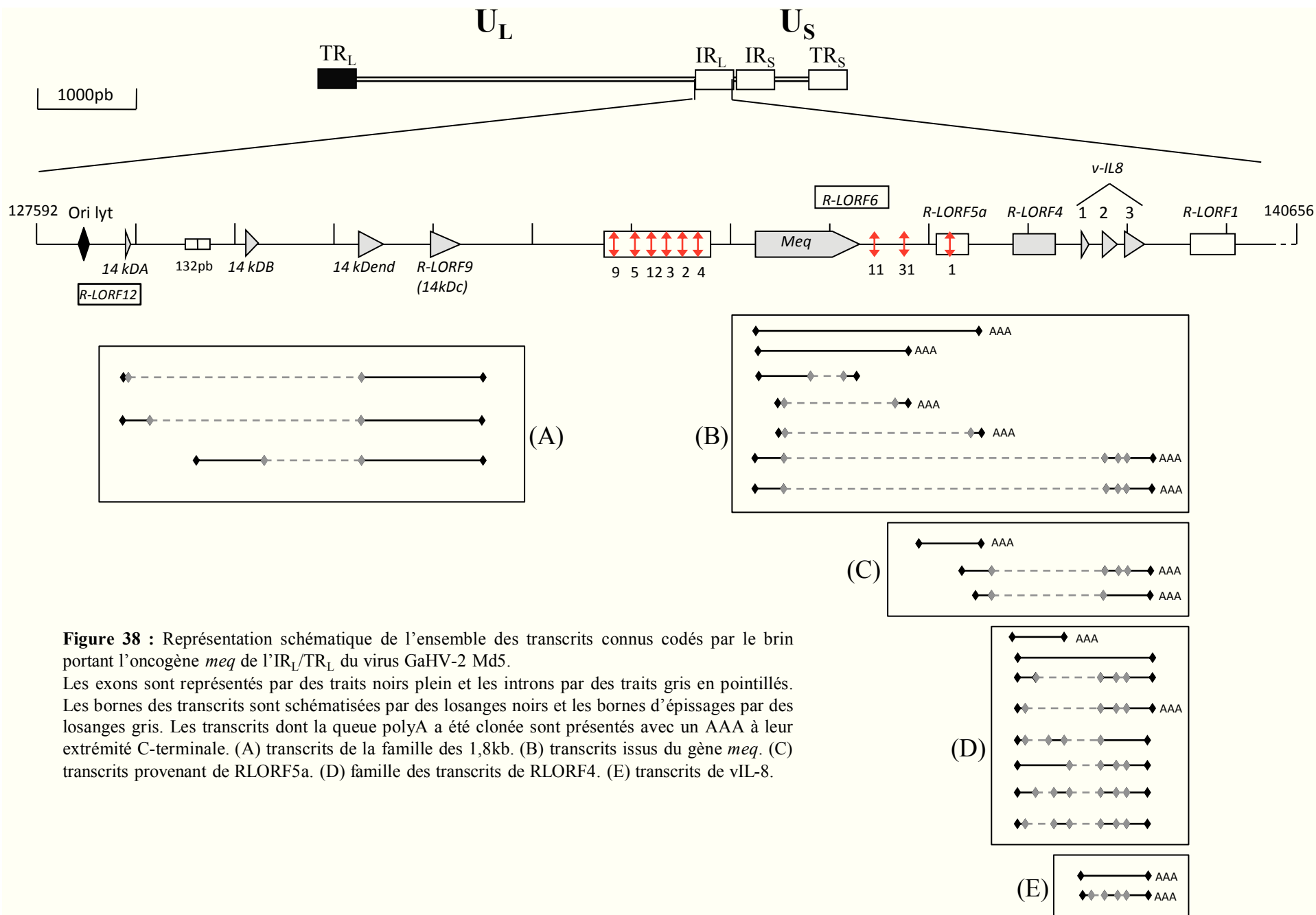
3.2.5.2.3.1. Le gène *meq* (Marek EcoQ)

Le gène *meq* (Marek's disease EcoRI-fragment Q ou RLORF-7), considéré comme l'oncogène majeur de GaHV-2, reste le gène de ce virus le plus étudié à ce jour. Il code pour une phosphoprotéine nucléaire de 339 aa impliquée dans de nombreux processus de transformation cellulaire. Cette protéine est plus fortement exprimée dans les cellules tumorales induites par GaHV-2 et dans les lignées cellulaires transformées que pendant le cycle lytique (Liu *et al.*, 1998; Ross *et al.*, 1997). L'expression du gène *meq* dans les leucocytes sanguins de poulet est détectable *in vivo* dès le début de l'infection et augmente pour être maximale au moment de l'apparition et du développement des lymphomes (Debba-Pavard *et al.*, 2008a). La protéine Meq possède de nombreux domaines caractéristiques de liaison à l'ADN, à l'ARN, un domaine de type Leucine-Zipper (domaine bZIP) et des séquences de ciblage vers le noyau et les nucléoles (Figure 37). La région C-terminale riche en proline (Figure 37), homologue à la protéine suppresseur de tumeur WT-1 (Wilms Tumor-1), comporte un motif de liaison à l'ARN ainsi qu'un domaine transactivateur (Qian *et al.*, 1995). Le domaine bZIP localisé dans la région N-terminale de la protéine est homologue à celui des oncogènes de la famille Fos/Jun (Jones *et al.*, 1992). Ce domaine permet à Meq de se lier à l'ADN, de s'homodimériser, ou de se dimériser avec d'autres protéines possédant généralement un domaine bZIP. Ainsi, Meq semble pouvoir former des hétérodimères avec de nombreux facteurs cellulaires, dont c-Jun (Qian *et al.*, 1995), junB, Fos, CREB, ATF1, ATF2, ATF3 (Levy *et al.*, 2003b), ICP4, p53, RB (Liu & Kung, 2000), CtBP, c-Myc (Osterrieder *et al.*, 2006) et Hsp70 (Zhao *et al.*, 2009a), avec cependant une affinité supérieure pour c-Jun (Levy *et al.*, 2003b). Par ailleurs, les hétérodimères Meq/Jun et les homodimères Meq/Meq se fixent sur des séquences nucléotidiques appelées MERE I et MERE II (Meq Responsive Elements). Le site MERE I (GAGTGATGACGTCATC) est un site de type AP-1 plutôt spécifique de l'hétérodimère Meq/Jun. Il est retrouvé au niveau des promoteurs des gènes viraux *meq* et *ICP4* ou cellulaires *IL-2* et *CD30* avec un effet cis-activateur associé (Burgess *et al.*, 2004; Levy *et al.*, 2003b). Le site MERE II (RACACACAY) spécifique de l'homodimère Meq/Meq, identifié près de l'origine de réplication de GaHV-2 conduit à l'inhibition de la transcription des gènes de la famille des 1,8 kb et codant pour pp38 (Levy *et al.*, 2003b). Ainsi, Meq pourrait réguler sa propre expression et inhiber celle des gènes intervenant dans la réplication virale, suggérant un rôle de ce gène dans l'établissement de la latence

ainsi que dans la transformation cellulaire et la lymphomagenèse. Enfin, des études récentes utilisant des virus recombinants a permis de montrer que l'homodimérisation de la protéine était nécessaire à l'induction de lymphomes (Brown *et al.*, 2009), mais qu'elle n'était pas suffisante pour induire la transformation à elle seule (Suchodolski *et al.*, 2009).

Plusieurs études ont été conduites pour rechercher l'impact de Meq sur le cycle cellulaire et sur les voies de régulation de l'apoptose. En premier lieu, il a été montré que lors d'expressions ectopiques dans les lignées hétérologues Rat-2, Meq co-localisait avec CDK2 (Cyclin Dependant Kinase 2) pendant les phases G1/S et S du cycle cellulaire, à la périphérie des nucléoles et dans les corps de Cajal (Liu *et al.*, 1999b). Il faut noter qu'un impact sur la formation du complexe RB/E2F et donc sur le cycle cellulaire n'a pas été mis en évidence à ce jour, mais il semblerait plutôt que CDK2 en phosphorylant Meq réduirait son activité de liaison à l'ADN et favoriserait son exportation vers le cytoplasme (Liu *et al.*, 1999b). Par ailleurs, Meq semble présenter un effet anti-apoptotique évident. En effet, lors d'expression ectopique dans la lignée aviaire DF1, Meq augmenterait la transcription de gènes tels que *Bcl-2*, *c-jun* et *c-ski* et diminuerait celle des gènes pro-apoptotiques *Fas* et *DAP5* (Levy *et al.*, 2005). De plus, Meq aurait également un effet anti-apoptotique par blocage du facteur TNF α (Tumor Necrosis Factor α) (Liu *et al.*, 1998). Enfin, un motif PLDLS (Pro-Leu-Asp-Leu-Ser) localisé dans la partie N-Terminale de la protéine Meq a été montré fonctionnel par établissement d'un lien direct entre la mutagenèse de ce motif et la perte du pouvoir oncogène du virus mutant correspondant (Brown *et al.*, 2006). Ce motif de type PXDLS est spécifique de l'interaction avec la protéine CtBP (C-terminal Binding Protein), qui est un co-répresseur transcriptionnel impliqué dans la régulation de la prolifération cellulaire ainsi que dans l'apoptose, mais dont le mécanisme d'action précis est encore peu connu. Il faut noter que des motifs PXDLS fonctionnels ont également été caractérisés chez d'autres oncogènes viraux essentiels, tels que E1A d'adénovirus ou l'antigène EBNA3C (Epstein Barr Nuclear Antigen 3C) (Brown *et al.*, 2006). Suite à ces résultats, il a été montré que des mutants de GaHV2 invalidés pour le gène *meq* étaient capables de se répliquer *in vivo* et *in vitro* mais n'exprimaient plus de pouvoir pathogène chez l'animal, faisant de ces mutants des candidats efficaces pour la vaccination en conditions expérimentales (Lee *et al.*, 2010; Lee *et al.*, 2008).

Un polymorphisme de *meq*, consistant en la présence d'un nombre variable de séquences répétées riches en proline (PRR, Proline-Rich Repeats) a été mis en évidence suite au séquençage du gène de diverses souches de GaHV-2 (Chang *et al.*, 2002a; Chang *et al.*, 2002b). Ainsi, quatre formes alléliques du gène *meq* ont été mises en évidence après analyse de l'ADN extrait de CEF infectés *in vitro*, la forme longue L-*meq* (Long-*meq* ; insertion de 180pb), la forme classique *meq*, la forme courte S-*meq*, (Short-*meq* ; délétion de 123pb) et la forme très courte VS-*meq* (Very Short-*meq* ; délétion de 276 pb) (Chang *et al.*, 2002b). Cependant, les auteurs ont seulement retrouvé les formes L-*meq* et *meq* *in vivo* (Chang *et al.*, 2002a), avec une majorité de la forme L-*meq* possédant neuf copies de PRR dans le génome des souches faiblement oncogénique (JM) ou non oncogénique (CVI988) durant toutes les



phases virales, et une majorité de l'isoforme *meq* (6 copies de PRR) durant les phases cytotolytiques et tumorales des souches virulente Md5 et RB-1B (Chang *et al.*, 2002a).

L'activité transactivatrice de L-Meq est moins importante que celle de Meq, suggérant un lien potentiel entre l'accroissement du nombre de PRR et une atténuation du pouvoir transactivateur de Meq (Chang *et al.*, 2002c). De plus, L-Meq semble inhiber la transactivation induite par Meq et réduire la réplication *in vitro* de la souche vvMd5 (Chang *et al.*, 2002c). Ainsi, L-Meq, qui est majoritaire durant la phase de latence des souches oncogéniques (Chang *et al.*, 2002a) ainsi que dans les cellules MSB1 (Chang *et al.*, 2002b), pourrait être impliquée dans le maintien de la latence des souches virulentes de GaHV-2.

L'analyse comparative des séquences de Meq a également permis d'identifier des mutations dans les motifs PRR des souches de GaHV-2 et une corrélation a été établie entre la présence de ces mutations sur les sites de transactivation ou de transrépression et la virulence de ces souches (Shamblin *et al.*, 2004). Il faut noter que le gène *meq* de la souche CVI988 présente également une délétion des trois nt du codon 193 (Kang *et al.*, 2007).

Par ailleurs, différentes études montrent qu'il existe plusieurs transcrits alternatifs de *meq* parmi lesquels, Meq/vIL-8 dont la région 3' est remplacée par les 2^{ème} et le 3^{ème} exons de *vIL-8* (Peng *et al.*, 1995), ΔMeq issu d'un épissage interne à *meq* (Okada *et al.*, 2007), Meq/RLORF5a, et un dernier transcrit dont le second exon commence entre *meq* et *RLORF5a* (Jarosinski & Schat, 2006) (Figure 38B), aboutissant à la production de différentes isoformes de Meq dont le rôle n'est pas encore clairement établi. Cependant il semblerait que certaines formes alternatives de la protéine Meq correspondant à ces épissages différentiels pourraient jouer un rôle dans la régulation négative de la latence et de la transformation induite par Meq. En effet la protéine Meq/vIL-8 qui a perdu la région transactivatrice terminale au profit de vIL8, a été observée à la fois pendant la phase lytique et la latence (Peng & Shirazi, 1996a). Elle se localise dans les mêmes compartiments nucléaires que Meq et semble se dimériser également avec c-Jun et se fixer au niveau des sites AP-1 (Anobile *et al.*, 2006; Peng & Shirazi, 1996a). Aussi, il est possible de lui supposer un rôle d'inhibiteur de la transactivation par Meq, par simple compétition avec la protéine native lors de l'assemblage des dimères transactivateurs. Par ailleurs, Okada et ses collaborateurs ont montré une forte expression du transcrit ΔMeq durant l'apoptose des lignées cellulaires transformées par GaHV-2 (Okada *et al.*, 2007). La protéine ΔMeq correspondante qui ne présente que les 100 premiers aa de la protéine Meq native, semblerait également avoir un effet négatif sur la transactivation induite par Meq et L-Meq (Okada *et al.*, 2007).

3.2.5.2.3.2. pp38/pp24

Les gènes précoces pp38 (R-LORF14a) et pp24 (R-LORF14), spécifiques de GaHV-2, sont localisés respectivement au niveau des jonctions entre l'U_L et l'IR_L ou l'U_L et le TR_L. De par leur

localisation, ces protéines présentent une région NH2 terminale identique puisque codée par les régions répétées et elles diffèrent dans leur partie COOH terminale codée par les 2 extrémités différentes de l'U_L. De ce fait, ces deux polypeptides possèdent un promoteur identique qui est caractérisé par la présence de l'origine de réplication lytique (Figure 28) (Katsumata *et al.*, 1998). Comme évoqué précédemment, ce promoteur bidirectionnel régulé négativement par l'homodimère Meq/Meq (Levy *et al.*, 2003b) et positivement par l'hétéropolymère pp38/pp24 (Ding *et al.*, 2008; Ding *et al.*, 2007) montre une expression plus importante des gènes de la famille des 1,8 kb en comparaison des gènes pp38/pp24 (Chen *et al.*, 2009; Shigekane *et al.*, 1999). Des homologues de ces gènes ont été retrouvés chez GaHV-1 (Izumiya *et al.*, 2001; Ono *et al.*, 1994), et seul un homologue tronqué de pp38 a été identifié chez HVT (Afonso *et al.*, 2001; Smith *et al.*, 1995).

L'expression des phosphoprotéines pp24 et pp38 peut être augmentée par un traitement à l'iododéoxyuridine (IUdR) ou par transfection du gène ICP4 (Pratt *et al.*, 1994) conduisant plutôt à attribuer un rôle pour ces protéines dans la réactivation et le cycle lytique. Bien que le rôle de *pp24* n'a pas été étudié, l'implication de *pp38* durant l'infection cytotytique précoce mais également durant la transformation a été observée.

La phosphoprotéine de 290 aa pp38, a été initialement identifiée comme un antigène associé à la transformation, exprimé dans le cytoplasme des cellules tumorales et des lignées lymphoblastoïdes (Cui *et al.*, 1990; Nakajima *et al.*, 1987). De plus, l'inhibition de la prolifération de lignées cellulaires lymphoblastoïdes transformées par des oligonucléotides « anti-sens » complémentaires de la région d'initiation de la traduction du gène *pp38* semble également suggérer un rôle de cette protéine dans le maintien du phénotype transformé (Xie *et al.*, 1996). L'utilisation d'un cosmide recombinant délété de *pp38* a permis de confirmer se rôle mais également de mettre en évidence le rôle prépondérant de *pp38* lors de l'infection cytotytique précoce des lymphocytes B (Gimeno *et al.*, 2005; Reddy *et al.*, 2002). De plus, il a été observé que l'utilisation de ce virus recombinant n'affecte pas la dissémination horizontale du virus, ni la réplication au sein des follicules plumeux et *in vitro* en DEF (Duck Embryo Fibroblasts) (Gimeno *et al.*, 2005; Reddy *et al.*, 2002).

Cependant, le lien entre ces rôles et la forme longue de la protéine pp38 est remis en cause par la découverte de deux formes épissées de la protéine pp38 *in vivo* et dans différentes lignées cellulaires (Li *et al.*, 2006; Piepenbrink *et al.*, 2009). Enfin, seul un rôle spécifique à la forme longue de pp38 a pu être identifié, conduisant à la surexpression de la succinate déshydrogénase mitochondriale, une enzyme pouvant potentiellement interférer avec la voie apoptotique mitochondriale (Piepenbrink *et al.*, 2009).

3.2.5.2.3.3. La famille des gènes de 1,8 kd

La famille des gènes 1,8 kb regroupe plusieurs transcrits très précoces spécifiques à GaHV-2 (Bradley *et al.*, 1989a; Peng *et al.*, 1992), au sein desquels deux protéines (14kDa ou 14kDb) ont été

identifiées (Figure 38A) dans le cytoplasme des cellules infectées par les souches oncogéniques et atténuées et dans les lignées transformées par GaHV-2 (Hong *et al.*, 1995). Elles proviennent de transcrits épissés, constitués de deux exons. Un premier exon, spécifique de chaque protéine 14kDa ou 14kDb, et un second commun aux deux messagers le 14kDend (Hong & Coussens, 1994). Le transcrit 14kDa, qui s'initie au niveau du promoteur bi-directionnel, encode pour une protéine de 85 aa (Hong & Coussens, 1994), tandis que le transcrit 14kDb, qui s'initierait en amont des répétitions de 132 pb, encode pour une protéine de 93 aa (Hong & Coussens, 1994; Tahiri-Alaoui *et al.*, 2009a). Ces deux phosphoprotéines, possédant un poids moléculaire respectif de 9,3 et 8,4 kDa, contiennent de nombreux sites potentiels de phosphorylation et de glycosylation aboutissant à la masse moléculaire apparente d'environ 14kDa (pp14).

De plus, deux IRES (Internal Ribosome Entry Site) ont été recensés sur le ARNm de 14 kDb. Le premier, situé en amont du premier exon de 14kDb (Tahiri-Alaoui *et al.*, 2009a), et le second, intersitronique, entre 14kDend et R-LORF9 (Tahiri-Alaoui *et al.*, 2009b). Ces deux IRES sont indépendamment fonctionnels, mais semblent pouvoir agir en synergie afin de permettre la traduction des deux protéines 14kDb et pp12 (Tahiri-Alaoui *et al.*, 2009a). Il faut noter que cette protéine pp12 de 107 aa et de masse moléculaire apparente de 12kDa peut être également traduite par le même mécanisme à partir du messenger de 14kDa.

Le rôle de ces trois protéines n'est pas clairement établi, néanmoins, une inhibition de la réplication de l'ADN et de la prolifération de cellules lymphoblastoïdes transformées par le GaHV-2 a été observée suite à l'utilisation d'oligonucléotides « anti-sens » complémentaires de cette famille d'ARN (Hayashi *et al.*, 1999; Kawamura *et al.*, 1991). Plus précisément, elles interagiraient avec le cycle cellulaire en favorisant le passage de la phase G1 à la phase S, ou en ayant un rôle positif lors des premières étapes de la réplication de l'ADN (Hayashi *et al.*, 1999).

3.2.5.2.3.4. Le gène R-LORF5a

Le gène *R-LORF5a* correspond à une ORF située entre *meq* et *R-LORF4* qui coderait pour une protéine potentielle de 107 aa (Figure 38) (Ohashi *et al.*, 1994), ainsi qu'un des miARN de GaHV-2, mdv1-miR-M1 (Figure 33) (Burnside *et al.*, 2006). Le transcrit de 0,6kb recouvrant cette ORF semble être modérément exprimé lors d'infection lytique de CEF par la souche vvRB-1B et très fortement dans des cellules lymphoblastoïdes transformées CU41 (Figure 38C) (Ohashi *et al.*, 1994). De plus, à l'instar de *meq* et de *R-LORF4*, des transcrits alternatifs issus de *R-LORF5a* montrent un épissage avec les exons 2 et 3 de *vIL-8* (Figure 38C) (Jarosinski & Schat, 2006). Cependant, des infections *in vitro* et *in vivo* avec un virus CVI988 ou RB-1B délété pour *R-LORF5a* montre un pouvoir infectieux comparable à celui des souches sauvages (Jarosinski *et al.*, 2005; Schat *et al.*, 1998), ce qui semblerait montrer que mdv1-miR-M1 ainsi que la protéine potentielle R-LORF5a ne seraient pas essentiels pour la réplication de GaHV-2.

3.2.5.2.3.5. Le gène R-LORF4

Un transcrit propre à R-LORF4 a été identifié durant la latence dans les lignées cellulaires MSB-1 issues de tumeurs et durant les cycles réplicatifs dans les CEF infectées par des souches caractéristiques de chaque niveau de virulence (Jarosinski *et al.*, 2003). R-LORF4 coderait potentiellement pour une protéine de 142 aa, qui n'a jamais été visualisée. Cependant deux peptides codés par R-LORF4 ont été retrouvés lors d'une étude protéomique par spectrométrie de masse tendant à valider l'expression réelle d'une protéine correspondant à cet ORF (Liu *et al.*, 2006). De plus, de nombreux transcrits alternatifs comprenant des phénomènes d'épissage interne à *R-LORF4* ainsi qu'avec les exons 2 et 3 de *vIL-8* (Jarosinski *et al.*, 2003; Jarosinski & Schat, 2006) ayant été retrouvés (Figure 38D), il est difficile de connaître par prédiction uniquement quelle est la protéine réellement produite à partir de R-LORF4. Par contre, R-LORF4 est important pour GaHV-2, puisque la délétion des deux copies du gène *R-LORF4* d'un virus vvRB-1B mène à une atténuation de la souche *in vitro* et *in vivo* (Jarosinski *et al.*, 2005) et que l'atténuation d'une souche vv+ (648a) après cent passages en CEF entraîne la délétion de tout ou partie de *R-LORF4* (Spatz, 2010).

3.2.5.2.3.6. Le gène vIL-8

Le gène *vIL-8*, codé par *R-LORF2*, a initialement été identifié suite à l'étude des produits épissés de *meq* (Figure 38) (Peng & Shirazi, 1996a). Le gène *vIL-8* s'exprime durant la phase lytique et est composé de trois exons codant pour une chemokine virale de type CXC homologue de l'*IL-8* aviaire (Liu *et al.*, 1999a; Parcells *et al.*, 2001). Le premier exon de *vIL-8* code pour le peptide signal tandis que les exons 2 et 3 composent la protéine mature qui possède toutes les grandes caractéristiques d'une chemokine (Liu *et al.*, 1999a). La seule différence consiste dans le remplacement de la séquence ELR (Glu-Leu-Arg) très conservée dans toutes les molécules IL-8-like cellulaires (Liu *et al.*, 1999a) par un motif DKR (Asp-Lys-Arg) qui cependant, semble aussi fonctionnel que le motif ELR lors de la pathogenèse de GaHV-2 (Cui *et al.*, 2004).

Le rôle de *vIL-8* dans la pathogenèse de GaHV-2 a été mis en évidence par l'observation d'une diminution significative de la réplication virale *in vivo* et d'une baisse de l'incidence tumorale induite par la délétion du gène *vIL-8* de la souche virale vvRB-1B (Parcells *et al.*, 2001). Plus précisément, *vIL-8* n'est pas nécessaire à l'établissement de la latence et sa délétion n'influe pas sur la réplication au niveau des follicules plumeux (Cui *et al.*, 2004). De plus, une expression plus importante de *vIL-8* a été observée lors d'infection *in vivo* par des souches vv et vv+ en comparaison des souches moyennement virulentes et avirulentes (Jarosinski *et al.*, 2003). Ainsi la protéine vIL-8 semble impliquée dans la phase précoce d'infection lytique et particulièrement dans le recrutement des lymphocytes T, à la suite de l'expression des récepteurs à cette chemokine activés par l'interféron γ . Cette caractéristique pourrait permettre à vIL-8 de jouer un rôle clef dans l'infection secondaire des

lymphocytes T par GaHV-2, suite à l'infection initiale des lymphocytes B (Schat & Xing, 2000). De plus, vIL-8 est capable d'être attractive pour les monocytes du sang circulant, et de ce fait, cette protéine pourrait participer à un mécanisme d'échappement de GaHV-2 à la réponse immune, tel que cela a déjà été décrit pour d'autres herpesvirus (Alcami & Koszinowski, 2000).

Comme cela a été décrit dans les paragraphes précédents un épissage différentiel a été mis en évidence pour vIL-8 (Figure 38), et différents résultats contradictoires posent question quant à la nécessité de ce peptide signal codé par l'exon 1. En effet, de multiples transcrits s'épissant avec l'exon 2 et l'exon 3 de vIL-8, excluant le premier exon de vIL-8 ont été identifiés (Figure 38B 38C 38D) (Jarosinski & Schat, 2006). De plus, à l'instar de ceux de vIL-8, les transcrits R-LORF5a/vIL-8 et R-LORF4/vIL-8 semblent être exprimés durant la phase précoce de réplication (Jarosinski & Schat, 2006). Par contre, la délétion du premier exon de vIL-8 chez un virus vvRB-1B entraîne une diminution de l'infection lytique précoce (Jarosinski & Schat, 2006). Cependant, il est difficile de conclure car avec ce virus délété de l'exon 1 de vIL-8 les transcrits R-LORF5a/vIL-8 et R-LORF4/vIL-8 n'étaient pas décelés lors d'infection *in vivo* (Figure 38B 38C 38D) (Jarosinski & Schat, 2006).

3.2.5.2.3.7. vTR

Le premier ARN fonctionnel découvert chez GaHV-2 correspond à une sous-unité ARN de la télomérase vTR, qui est le seul homologue viral de TR (Telomerase RNA) identifié à ce jour (Figure 39) (Fragnet *et al.*, 2003). La télomérase est un complexe ribonucléoprotéique qui permet l'élongation des télomères localisés à l'extrémité des chromosomes dont ils assurent l'intégrité lors des divisions cellulaires. Le complexe télomérasique est composé d'une protéine nommée TERT (Telomerase Reverse Transcriptase) qui possède une activité transcriptase inverse et un ARN (TR) comportant une séquence matricielle (CUAACC) complémentaire des motifs télomériques (GGTTAG), retrotranscrite par TERT. La télomérase semble impliquée dans environ 85 % des cancers humains et l'activité de cette enzyme semble essentielle à la prolifération des cellules tumorales et à leur immortalisation.

vTR est spécifique de GaHV-2, il est codé à l'extrémité des répétitions TR_L et IR_L, sur le brin opposé à l'oncogène *meq*, proche de la région comportant des séquences télomériques caractéristique de la jonction IR_L/IR_S ou TR_L/TR_S (Figure 27). vTR présente une forte homologie de séquence (88%) avec son homologue aviaire, chTR (Figure 39), tout en étant plus efficace que ce dernier pour reconstituer une activité télomérase avec chTERT en cellule (Fragnet *et al.*, 2005). Par ailleurs, contrairement au promoteur de chTR, le promoteur de vTR semblerait essentiellement piloté par le proto-oncogène cellulaire c-Myc, qui lui conférerait une efficacité de transcription quatre fois supérieure à celle de son homologue cellulaire dans les cellules transformées MSB-1 (Shkreli *et al.*, 2007). De plus, l'importance de c-Myc dans la régulation de vTR et dans l'oncogenèse de GaHV-2 a

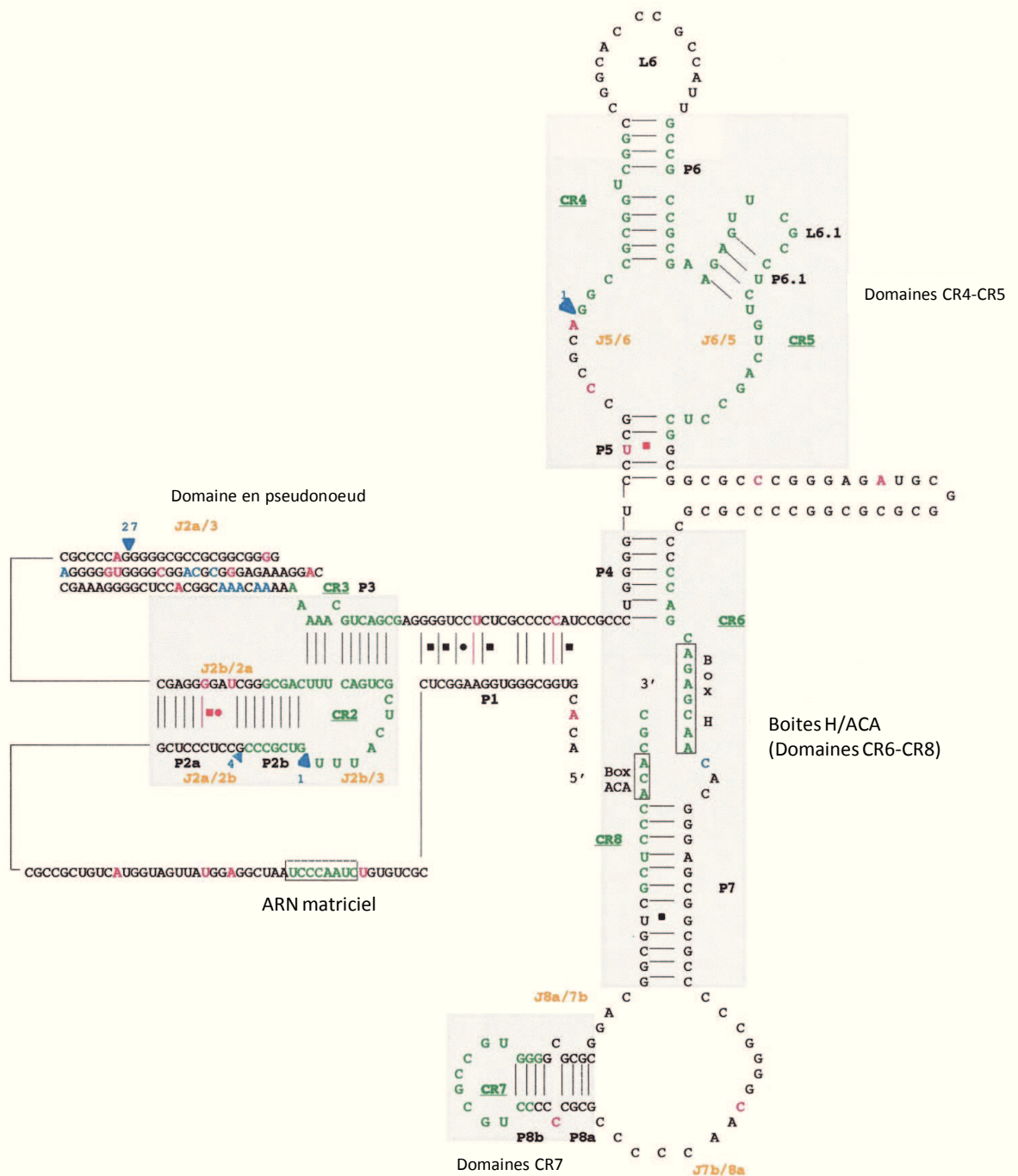


Figure 39 : Structure secondaire de vTR (d'après Fragnet et al., 2003)

Les régions appariées en hélice (P1 à P8) sont numérotées à partir de l'extrémité 5' de vTR. Les régions jonctionnelles (J) entre 2 hélices sont indiquées en orange et identifiées par rapport aux régions d'appariement bordantes. Les 4 domaines conservés dans les TR de vertébrés sont encadrés. La région matricielle (CR1) et les boîtes H et ACA sont délimitées par des rectangles noirs. Les mutations et les appariements additionnels dans la séquence de vTR par rapport à chTR sont indiqués en rouge. Les nucléotides additionnels dans la séquence de vTR sont indiqués en bleu et les régions de délétion sont représentées par des triangles bleus comportant le nombre de bases délétées.

été confirmée *in vivo* suite au remplacement du promoteur de vTR par celui de chTR dans la souche virale RB-1B (Chbab *et al.*, 2010).

Au vu du rôle prépondérant de la télomérase dans les processus d'immortalisation, la dérégulation de cette dernière par la présence de la sous-unité vTR semble être un facteur déterminant dans les processus de tumorigenèse viro-induite par GaHV-2. En effet, une forte activité télomérase a été détectée dans la lignée transformée MSB-1 (Djeraba-AitLounis *et al.*, 2004), et une augmentation d'activité télomérase a été mise en évidence *in vivo* dans les PBL de poulets infectés par une souche vvRB-1B, parallèlement à une forte expression de vTR observée à partir du 7^{ème} jour post infection et perdurant jusqu'à la mort des animaux (Shkreli *et al.*, 2007). De plus, la délétion des deux copies de vTR dans le génome de la souche pathogène vvRB-1B réduit de 50% l'incidence tumorale et diminue la dissémination des lymphomes, en comparaison du virus parental (Trapp *et al.*, 2006). Il est également intéressant de noter que la seule mutation qui semble influencer l'efficacité de vTR a été identifiée dans certains isolats de la souche vaccinale CVI988 (Debba-Pavard *et al.*, 2008b; Fragnet *et al.*, 2005; Spatz & Silva, 2007a). Cette mutation A en G touche les boîtes H/ACA qui sont impliqués dans la localisation nucléaire (nucléoles, corps de cajal) et induisent une diminution de l'efficacité de vTR.

3.2.5.2.3.8. Autres ORF présentes dans l'IR_L/TR_L.

L'expression de la plupart des ORF présentes sur le brin complémentaire de l'oncogène *meq* au sein des régions répétées IR_L et TR_L n'a pas été validée pour le moment. Parmi ces ORF, seuls des peptides issus de *R-LORF10*, *R-LORF11* ainsi que la protéine 23kD ont été détectées (Figure 40) (Liu *et al.*, 2006; Peng & Shirazi, 1996b). Bien que les protéines de *R-LORF8*, *R-LORF12* et *R-LORF13* n'aient pas été retrouvées dans un contexte d'infection par GaHV-2, des interactions *in vitro* avec la protéine cellulaire C1QBP (C1q Binding Protein) (Niikura *et al.*, 2004), également appelée SF2P32 (p32 subunit of splicing factor ASF/SF2), ont été décrites pour les protéines codées par *R-LORF8* et *R-LORF13*. C1QBP est impliquée dans la régulation de l'épissage par inhibition de la liaison à l'ARN et par phosphorylation de la protéine SR ASF/SF2 (Petersen-Mahrt *et al.*, 1999). Il est intéressant de noter que la protéine ICP27 d'HHV-1 se lie également à la protéine C1QBP (Bryant *et al.*, 2000).

La protéine potentielle codée par *R-LORF12* semble interagir avec la protéine cellulaire p23 (Niikura *et al.*, 2004), également appelée TCTP (Translationally Controlled Tumor Protein) encodée par le gène *TPT1* (Tumor Protein Translationally controlled 1). Cette protéine ubiquitaire est très conservée chez l'ensemble des eucaryotes. *TPT1* possède de nombreuses fonctions dont des rôles dans la croissance cellulaire, le cycle cellulaire, la transformation, l'apoptose mais également un rôle dans l'activation de PKR (Protein Kinase R ou double strand RNA Dependent Protein Kinase) généralement impliquée dans la défense antivirale.

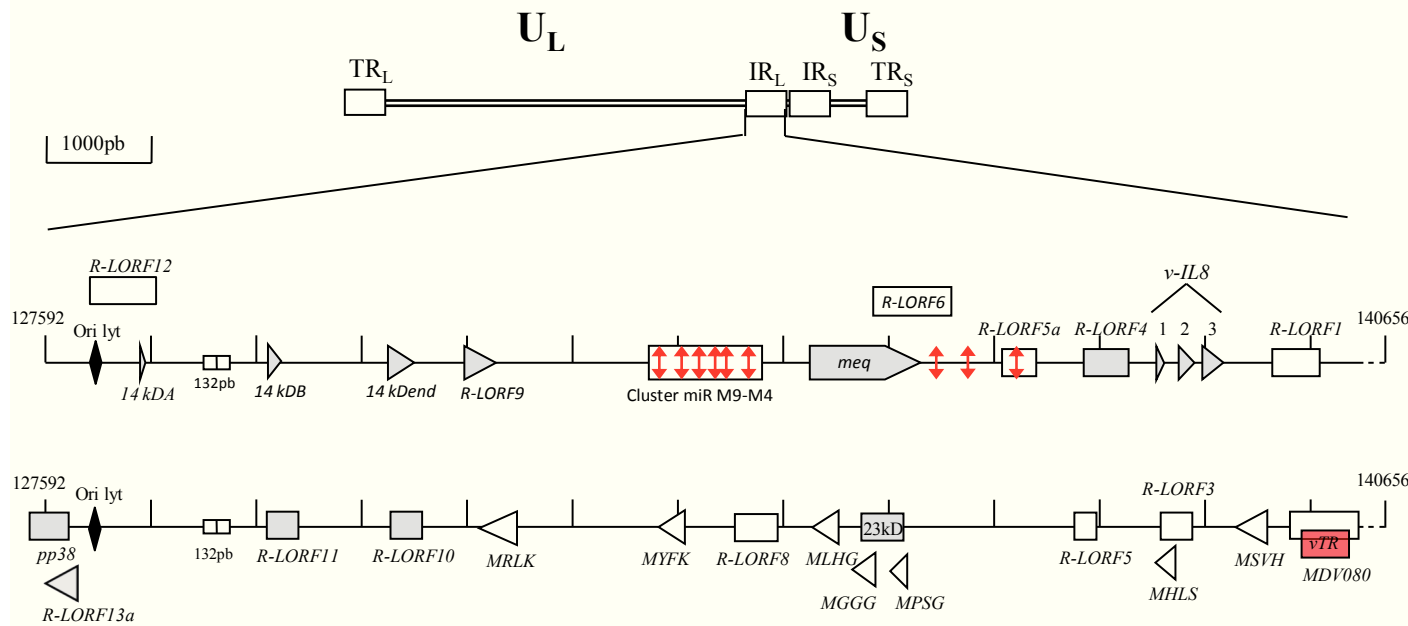


Figure 40 : Représentation schématique des deux brins codants de l'IRL du génome de GaHV-2 Md5.

Les ORF dont les protéines associées ont été retrouvées sont représentés en gris. Les ORF dont aucune protéine n'a été identifiée à ce jour sont représentés en blanc. Les ARN fonctionnels (microARN et vTR) sont représentés en rouge. L'origine de réplication lytique du virus (Ori lyt) est également positionnée (losange noir). Les positions correspondent au génome GaHV-2 md5.

La protéine 23kD, localisée sur le brin complémentaire de l'oncogène *meq* (Figure 40), a été détectée dans les fractions nucléaires de cellules infectées par GaHV-2 (Peng & Shirazi, 1996b). Cette protéine interagit avec l' α -enolase (ENO1) (Niikura *et al.*, 2004), qui est retrouvée à la surface des lymphocytes B et T, et dont l'expression semble corrélée à la progression des tumeurs.

La protéine codée par *R-LORF10* interagit *in vitro* avec la chaîne invariante Ii du CMH de classe II (CMH II) (Niikura *et al.*, 2004). Contrairement aux autres herpesvirus qui ont tendance à diminuer l'expression du CMH II, une surexpression de ce complexe a été identifiée *in vitro* et *in vivo* par GaHV-2 (Niikura *et al.*, 2007). De plus, un virus délété de *R-LORF10* montre une virulence atténuée *in vivo*, tandis qu'aucune différence n'est détectée *in vitro* bien qu'une diminution du CMH II à la surface de ces cellules soit observée (Kim *et al.*, 2010).

4. Objectifs de travail

Les miARN apparaissent comme les régulateurs principaux de l'expression génique pour l'ensemble des eucaryotes. Ils sont également essentiels lors des infections virales quelque soit le type de virus et surtout pour les virus encodant des miARN tels que les herpesvirus. Bien que tous les herpesvirus n'encodent pas des miARN, ils représentent la famille de virus encodant la majorité des miARN viraux identifiés. Les miARN des herpesvirus semblent impliqués dans de nombreux processus, tels que la régulation du cycle viral, la dérégulation de la réponse immunitaire cellulaire et/ou l'oncogenèse. Chez GaHV-2, 25 miARN matures ont été identifiés (Burnside *et al.*, 2006; Burnside *et al.*, 2008; Morgan *et al.*, 2008). Ils sont répartis au sein des régions répétées TR_L/IR_L et TR_S/IR_S du virus, contenant les gènes spécifiques de GaHV-2.

Suite à l'établissement de banques de miARN de GaHV-2 RB-1B *in vivo*, le laboratoire TLVI (Transcription Lymphome Viro-Induit) s'est intéressé aux mécanismes de régulation de ces derniers ainsi qu'à leurs rôles dans la lymphomagenèse induite par GaHV-2. Une partie de l'équipe s'est intéressée aux miARN des régions R_S et j'ai eu en charge l'étude des miARN des régions R_L. Ainsi, mon travail de thèse a consisté en l'étude des différents mécanismes transcriptionnels et post-transcriptionnels liés aux régions IR_L/TR_L. De plus, suite aux données des banques de miARN montrant une grande variation d'expression des miARN viraux entre eux et au sein d'un même cluster nous nous sommes demandé quels facteurs pouvaient influencer leur régulation et plus particulièrement leur transcription. Pour répondre à cette problématique, différentes approches ont été suivies. Tout d'abord, une recherche et analyse du promoteur responsable de leur transcription, mais également, au vu de la complexité de transcription de la région TR_L/IR_L, une recherche plus globale des transcrits pouvant permettre leur transcription notamment lors de la phase de latence/oncogenèse du virus. Par ailleurs, suite à la découverte de l'homologie de séquence de la région « seed » entre mdv1-miR-M4-5P et le miARN cellulaire miR-155, au vu du rôle primordial de ce dernier dans la régulation des lymphocytes, la régulation de la réponse immunitaire et son impact dans l'oncogenèse, ainsi que de la prévalence de mdv1-miR-M4-5P au sein des différentes banques de miARN, nous avons comparé leurs actions sur différentes cibles cellulaires et virales, montrant un rôle potentiel dans l'oncogenèse et le maintien de la latence.

MATÉRIELS ET MÉTHODES

1. Lignées cellulaires et souches virales

1.1. Lignées cellulaires

Six lignées cellulaires de poulet ont été utilisées lors des différentes études. La lignée LMH établie à partir de cellules épithéliales de carcinome hépatique (Kawaguchi *et al.*, 1987) est cultivée en milieu DMEM : F12 (Lonza). Afin d'optimiser l'adhérence de ces cellules, les boîtes de culture sont pré-incubées 15 minutes (min) à 4°C avec de la gélatine (0,5%). La lignée DF-1 établie à partir de cellules épithéliales de poulet est cultivée en milieu DMEM (Lonza).

La lignée lymphocytaire T-CD4⁺ MSB-1 obtenue à partir de lymphomes de rate induits par la souche GaHV-2 BC-1 est également co-infecté par la souche GaHV-3 HPRS-24 (Akiyama & Kato, 1974), la lignée lymphocytaire T-CD4⁺ PA-9 obtenue à partir de lymphomes induits par la souche GaHV-2 HPRS-16 (Le Rouzic & Perbal, 1996), ainsi que la lignée lymphocytaire 54-O établie à partir de tumeurs ovariennes d'un poulet infecté par la souche GaHV2 RB1B reconstituée à partir d'un bacmide bac-RB1B (Muylkens *et al.*, 2010), possèdent le génome de GaHV-2 intégré au génome cellulaire. Ces lignées sont cultivées en milieu RPMI-1640 (Lonza), additionné de pyruvate de sodium (1mM). La lignée lymphoblastoïde B DT-40 (Baba *et al.*, 1985) dérive d'une tumeur induite par le virus leucosique aviaire RAV-1 (Rous-Associated Virus). Elle est cultivée en milieu DMEM (Lonza) additionné de pyruvate de sodium (1mM) ainsi que de tryptose broth phosphate (TBP ; 1,475 g/L).

Ces cellules sont cultivées en présence de 10% de sérum de veau foetal (SVF) et de 5% de sérum de poulet (SP) dans une atmosphère contenant 5% de CO₂. L'ensemble des lignées cellulaires sont incubées à 41°C.

1.2. Souches virales

La souche hypervirulente GaHV-2 RB-1B (vvRB-1B) maintenue sous forme d'une suspension de PBL récoltés à partir de poulets B¹³/B¹³ atteints de maladie de Marek 42 jours pi est conservée à raison de 10⁸ cellules par ml sous forme congelée dans l'azote liquide.

1.3. Préparation des CEF et infections virales

Les CEF primaires sont obtenus par trypsination d'embryons de poulet B¹³/B¹³ de 10 jours (Parker & Schierman, 1983) et sont mis en culture en milieu DMEM supplémenté de pyruvate de sodium (1mM), de TBP (1,475 g/L), de SVF (2,5%) et de SP (1,25%) à raison de 7.10⁶ cellules par flasque de 75cm². Ces cellules sont incubées à 41°C dans une atmosphère contenant 5% de CO₂.

Six heures avant l'infection, 4.10^5 de CEF secondaires sont mis en culture en plaque de 6 puits, puis infectées à raison de $1,5.10^6$ PBL par puits. Trois infections successives sont réalisées à partir d'un tiers des cellules du passage précédent, puis les cellules de chaque puits sont récupérées afin d'en extraire les ARN.

2. Les vecteurs de clonages et d'expression

Le plasmide pGEM-T Easy (Promega) permet le clonage direct de produits de PCR. Linéarisé par l'endonucléase EcoRV, il présente à ses extrémités une thymine capable de s'apparier à l'adénine ajoutée aux amplicons par certaines Taq polymérase. Ce plasmide de 3015 pb code pour la *beta*-lactamase et comporte le gène de la sous-unité *alpha* de la *beta*-galactosidase (cassette LacZ), dans lequel est localisé le site de clonage multiple encadré des promoteurs T7 et SP6, ce qui permet un criblage blanc-bleu des colonies possédant le plasmide recombinant (Figure 41A).

Le plasmide pcDNA3 (Invitrogen) permet l'expression de gènes clonés au niveau du site de clonage multiple en systèmes eucaryotes sous contrôle du promoteur du cytomégalovirus (Figure 41B).

Le plasmide pCDNAMLuc possède le gène codant la luciférase *Firefly* sous contrôle du promoteur du cytomégalovirus, permet l'expression de ce gène rapporteur en cellules eucaryotes.

Le plasmide pRL-TK (Promega), possédant le gène codant la luciférase *Renilla* sous contrôle du promoteur de la Thymidine Kinase (TK) du virus HHV-1, permet l'expression de ce gène rapporteur en cellules eucaryotes (Figure 41C).

Le vecteur d'expression pGL3 Basic (Promega) a été utilisé afin de tester l'efficacité de séquences promotrices clonées au niveau du site de clonage multiple situé en amont du gène rapporteur codant la luciférase *Firefly* (Figure 41D).

Le vecteur d'expression pcDNA3.1-V5-His-TOPO (Invitrogen) permet l'expression sous contrôle du promoteur du cytomégalovirus de protéines taguées par l'épitope V5 dérivé des protéines P et V du paramyxovirus SV5 (Figure 41E). De plus, ce vecteur linéarisé par l'exonucléase EcoRV, permet le clonage direct des produits de PCR.

Ces cinq plasmides codent pour la *beta*-lactamase, ce qui permet la sélection des bactéries transformées en milieu gélosé additionné d'ampicilline de concentration finale de 40 mg/L.

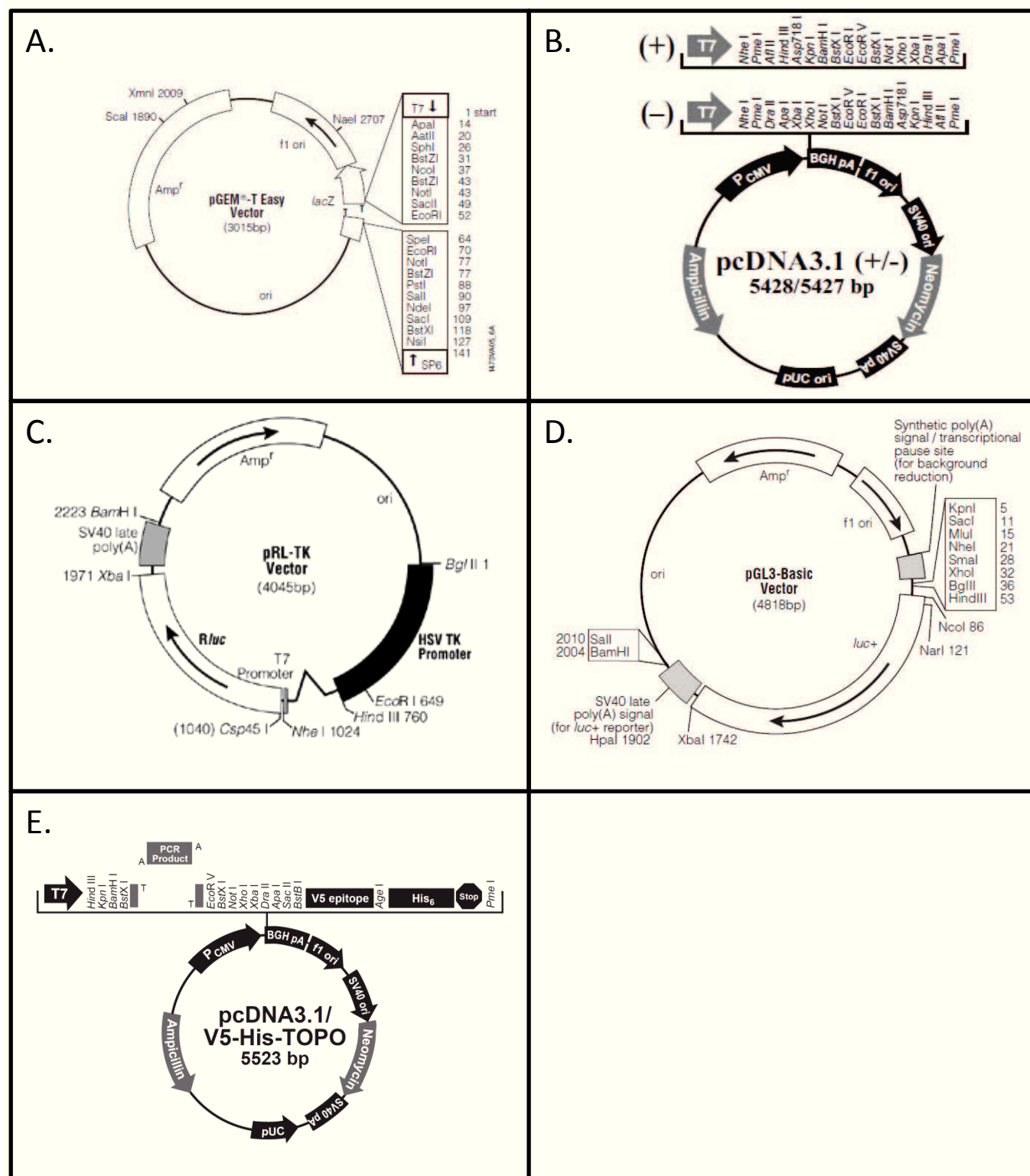


Figure 41 : Carte des vecteurs pGEMT-easy (A), pcDNA3.1 (B), pRL-TK (C), pGL3-basic (D) et pcDNA3.1-V5-His-TOPO (E).

3. Techniques générales de biologie moléculaire

3.1. Amplification génique

Les amplifications géniques sont effectuées dans un volume final de 50 µl, contenant 2,5 unités de Taq DNA polymérase (Promega), 100 µM de chaque amorce (Eurogentec), 200 µM de chaque désoxy-ribonucléotide, 10 mM de Tris-HCl pH 9, 50mM de KCl, 1,5 mM de MgCl₂, 0,1% de Triton X-100 et de 500 ng d'ADN matriciel.

Trente cycles d'amplification sont réalisés : 1 min à 94°C pour la dénaturation, 1 min à une température inférieure de 5°C à la température de fusion de l'amorce de plus faible T_m pour l'hybridation, et 1 min à 72°C pour la polymérisation. Un cycle supplémentaire de 10 min à 72°C est effectué afin de terminer la synthèse des brins.

3.2. Electrophorèse d'ADN

Les fragments d'ADN sont séparés par électrophorèse sur gel d'agarose (0,8%) en Tris-borate-EDTA (TBE) pour les gels analytique ou en Tris-acétate-EDTA (TAE) pour les gels préparatifs. La présence de 0,1 µg/ml de bromure d'éthidium dans les gels permet la visualisation de l'ADN sous rayonnement ultraviolet. Pour faciliter le dépôt, 1/5^{ème} de tampon de charge (0,3 % de bleu de bromophénol dans une solution de glycérol à 50 %, TBE 5X) est ajouté aux échantillons à analyser. La migration est effectuée sous un courant de 100V (TBE) ou 50V (TAE) avec le SmartLadder (Eurogentec) comme marqueur de taille.

3.3. Purification des acides nucléiques

Les fragments d'ADN sont purifiés à l'aide du kit NucleoSpin Extract II (Macherey-Nagel), fondé sur l'affinité de l'ADN pour la silice en présence de fortes concentrations en sel. Selon les recommandations du fournisseur, deux volumes de solution de solubilisation (NT) contenant des sels chaotropiques sont ajoutés pour un volume d'agarose excisé. Après incubation 5 min à 55°C, l'agarose solubilisé contenant le fragment d'ADN d'intérêt est déposé sur une colonne de silice et centrifugé 1 min à 11000g. Un lavage est ensuite réalisé par 700 µl de solution NT3 (composé de 4/5 d'éthanol

100%), suivie de deux centrifugations à 11000g pendant 1 et 2 min. L'ADN retenu dans la colonne est enfin élué par 18 µl d'eau ultra pure MilliQ lors d'une centrifugation de 1 min à 11000g.

3.4. Digestion enzymatique d'ADN

Les digestions d'ADN plasmidique ou des produits issus de PCR par les endonucléases de restriction (Promega) sont réalisées en utilisant 2 à 5 unités d'enzyme par µg d'ADN. Les réactions sont effectuées en présence du tampon de digestion spécifique des différentes enzymes utilisées. Les réactions sont incubées 1 heure à 37°C.

3.5. Déphosphorylation des vecteurs de clonages

Dix microgrammes de vecteur linéarisé sont ajoutés à un mélange réactionnel contenant 500mM de Tris-HCl, 1mM d'EDTA pH 8,5 et 10 unités de phosphatase alcaline (Promega). Après 20 min d'incubation à 37°C, l'enzyme est dénaturée par la chaleur pendant 20 min à 56°C. Après 5 minutes dans la glace, un second cycle est réalisé pendant 20 min à 37°C, puis 20 min à 65°C après ajout de 5 unités de phosphatase alcaline. Le vecteur déphosphorylé est purifié et concentré avant ligation soit en utilisant les colonnes Micropure et les membranes Microcon 100 (Millipore) soit par extraction par un volume d'un mélange phénol-chloroforme-alcool isoamylique (50-49-1) et précipitation par 3 volumes d'alcool à 100% en présence de 0,3 M d'acétate de Na.

3.6. Ligation vecteur/insert

Les réactions de ligation sont réalisées suivant un rapport molaire vecteur/insert de 1/3. Les réactions, réalisées dans un volume final de 10 µl contenant 3 unités de T4 DNA Ligase (Promega), 10mM d'ATP, 300mM de Tris-HCl, 100mM de MgCl₂, 100mM de DTT et 50 ng de vecteur, sont incubées 16 heures à 12°C. Lors des clonages dans le vecteur d'expression pGL3basic, les promoteurs sont insérés en amont du gène codant la luciférase Firefly, au niveau des sites de restriction *KpnI* /*HindIII* déphosphorylés en 5'. Pour les clonages dans le vecteur pRLTK, les inserts sont insérés en aval du gène codant la luciférase Renilla, au niveau des sites de restriction *XbaI* /*SpeI* ou *NotI* déphosphorylés en 5'.

3.7. Préparation des bactéries électro-compétentes

Une pré-culture en milieu LB est réalisée à 37°C pendant une nuit sous agitation. Dix millilitres de cette pré-culture sont ensuite utilisés pourensemencer 1L de milieu LB et incubés à 37°C jusqu'à obtention d'une $DO_{600nm}=0,5$. La croissance est alors arrêtée en plaçant la culture à 4°C. Après 10 min de centrifugation à 7000g à 4°C, cinq lavages sont effectués par 500 ml d'eau ultra pure à 4°C. Après un dernier lavage par 10 ml de glycérol 10% le culot bactérien resuspendu en glycérol 10%, puis réparti en fractions aliquotes de 50µl et conservé à -80°C.

3.8. Transformation bactérienne

Les bactéries *Escherichia coli* TG1 de génotype SupE hsdΔ5 thi Δ(lac-proAB)/F' [traD36 proAB+ lacIq lacZΔM15] sont classiquement utilisées, car elles permettent l'α-complémentation du gène de la β-galactosidase. Un microlitre de produit de ligation est ajouté à 50 µl de bactéries TG1 rendues "compétentes". L'entrée du plasmide dans les bactéries se fait par électroporation avec l'appareil Easyject (EquiBio), dans les conditions standardisées pour *E.coli* (2500 V pendant 5 ms avec une impédance de 25 µF). Immédiatement après électroporation, 150 µl de milieu de culture LB sont ajoutés aux bactéries. Après 30 min d'incubation à 37°C, celles-ci sont alors étalées sur du milieu LB/agar/ampicilline (0,1 g/L) lors d'un clonage dans les vecteurs pGL3, pcDNA3.1 et pRL-TK ou sur du milieu LB/agar/ampicilline (0,1 g/L) contenant 60 g/L de X-gal et 40 g/L d'IPTG lors des clonages dans le vecteur pGEM-T Easy, permettant un criblage blanc/bleu. Les boîtes de Pétri correspondantes sont finalement incubées 16 à 18 heures à 37°C.

3.9. Criblage des bactéries contenant le plasmide recombinant

Afin de sélectionner les clones bactériens comportant le plasmide recombinant, les colonies blanches sont repiquées sur LB/agar/ampicilline (0,1 g/L) contenant 60 g/L de X-gal et 40 g/L d'IPTG pour le stockage et mises en suspension dans le mélange de PCR pour le criblage des colonies renfermant un insert d'intérêt. Pour ce faire, chaque colonie bactérienne est directement ajoutée à 20 µl de mélange de PCR, contenant une unité de Taq DNA polymérase (Promega), 100 µM de chaque amorce de criblage (Eurogentec), 150 µM de chaque désoxy-ribonucléotide, 10 mM de Tris-HCl pH 9, 50 mM de KCl, 1,5 mM de MgCl₂ et 0,1 % de Triton X-100. Trente cycles d'amplification sont réalisés (94°C-1min, 55°C-1min, 72°C-1min). L'analyse des produits d'amplification est effectuée par électrophorèse en gel d'agarose/TBE.

3.10. Midi-préparation d'ADN plasmidique

Les clones d'intérêt sont mis en culture sous agitation dans 200 ml de milieu LB pendant 16h à 37°C. L'extraction de l'ADN s'effectue à l'aide du kit NucleoBond XTra Midi (Macherey-Nagel) par la méthode de lyse alcaline suivie d'une purification sur colonne de silice dans les conditions décrites par le fournisseur. La quantité d'ADN plasmidique extrait est estimée par mesure de l'absorbance à 260 nm sachant qu'une unité de densité optique correspond à 50 µg/ml d'ADN. Plusieurs digestions enzymatiques sont systématiquement réalisées afin de s'assurer de la qualité et de la spécificité de l'ADN plasmidique extrait.

Le séquençage des plasmides extraits (MWG) est systématiquement réalisé en utilisant les amorces adaptées au vecteur de clonage et des amorces internes si nécessaire (Tableau 5).

4. Elaboration des différentes constructions

4.1. Clonage des microARN

Les pré-miARN gga-pré-miR-155 et gga-pré-miR-21 ont été amplifiés par PCR à partir d'ADNg de la lignée cellulaire de poulet PA-12 au moyen des couples d'amorces ss-gga-pré-miR-155/rev-gga-pré-miR-155 et ss-gga-pré-miR-21/rev-gga-pré-miR-21 (Tableau 6). Les pré-miARN mdv1-pré-miR-M4 et mdv1-pré-miR-M7 ont été amplifiés à partir d'ADNg de GaHV-2 RB-1B avec les couples d'oligonucléotides ss-mdv1-pré-miR-M4/rev-mdv1-pré-miR-M4 et ss-mdv1-pré-miR-M7/rev-mdv1-pré-miR-M7 (Tableau 6). L'ensemble de ces pré-miARN ont été amplifiés avec des oligonucléotides contenant des extensions *NheI* et *HindIII*. Le produit de PCR digéré est ensuite inséré entre les sites *NheI* et *HindIII* du vecteur pcDNA3.1(+) afin de former les constructions pcDNA-gga-miR-155, pcDNA-gga-miR-21, pcDNA-mdv1-miR-M4 et pcDNA-mdv1-miR-M7.

4.2. Clonage des cibles virales et cellulaires des microARN

Un large fragment de la 3'UTR des gènes cellulaires *GPM6B*, *MAP3K7IP2*, *CEBPβ*, *RREB1*, *c-Myb* et *PU.1* contenant des miRE potentiels de mdv1-miR-M4-5P ont été amplifiés par PCR à l'aide des amorces « sens » (ss) et « anti-sens » (rev) correspondantes (Tableau 7). Ces amplicons ont été obtenus à partir d'ADNg de la lignée cellulaire PA-12 issue d'un poulet domestique de race Leghorn blanche. Les 500 nucléotides entourant les miRE des cibles virales UL-28 et UL-32 ont été amplifiées par PCR à partir d'ADNg de la souche GaHV-2 RB-1B avec respectivement les amorces « sens » (ss-UL-28 et ss-UL-32) et « anti-sens » (rev-UL-28 et rev-UL-32) (Tableau 7). Après digestion par les

oligonucléotides de séquençage	Séquence 5' → 3'
ss-pRL-TK	GCAAGAAGATGCACCTGATG
rev-pRL-TK	TCCTGTCATTCTAGTTGTGG
M13 sens (-21)	GTAAAACGACGGCCAGT
M13 reverse (-29)	CAGGAAACAGCTATGAC
ss-pGL3	CTAGCAAAATAGGCTGTCCCCAGTGC
rev-pGL3	GATAGAATGGCGCCGGGCCCTTTC
UL-28-mid*	CAATGAGAGACACTACTACCGA
UL-32-mid*	GGGTGATGCCAAAGTCTTGTA

Tableau 5 : Oligonucléotides utilisés pour le séquençage des banques et des différentes constructions.

* Oligonucléotides internes aux ORF UL-28 et UL-32.

Nom	Séquence 5' → 3'	Site de restriction
ss-gga-pré-miR-155	TTT <u>GCTAGC</u> AGAGTTCTTCTGTAGGCTGT	<i>NheI</i>
rev-gga-pré-miR-155	TTTAAGCTTCACAGCAGCAGTTGTTCCCT	<i>HindIII</i>
ss-gga-pré-miR-21	TTT <u>GCTAGC</u> TGTTGCCAGCCGTCGTGACAT	<i>NheI</i>
rev-gga-pré-miR-21	TTTAAGCTTGCTGGAGATGGGTGAGCAAA	<i>HindIII</i>
ss-mdv1-pré-miR-M4	TTT <u>GCTAGC</u> TGGGCTTGTTTGAATGTCC	<i>NheI</i>
rev-mdv1-pré-miR-M4	TTTAAGCTTTACGATGGTGTTCAGCGAT	<i>HindIII</i>
ss-mdv1-pré-miR-M7	TTT <u>GCTAGC</u> GATGCTCTCTAGCCAAGAGAG	<i>NheI</i>
rev-mdv1-pré-miR-M7	TTTAAGCTTGCGAGTTCTGAGGACACATTT	<i>HindIII</i>

Tableau 6 : Oligonucléotides utilisés pour l'élaboration des vecteurs d'expression des microARN.

Les sites de restriction servant au clonage des amplicons sont soulignés au sein de la séquence.

Nom	Séquence 5' → 3'	Site de restriction
ss-CEBPβ	CAGTCTAGA*CTTTAAGCATGGCTGGGAAT	<i>XbaI</i>
rev-CEBPβ	CAGACTAGTAAAGTTTATTAAAAATCGTTC	<i>SpeI</i>
ss-CEBPβ-mut	GCTACATTTGAAT TCGA §AATGAAGCATTTTAATAAAC	
rev-CEBPβ-mut	AAATCGTTCATT TCGA ATTCAAATGTAGCAAGAATGC	
ss-MAP3K7IP2	CAGTCTAGACACAGTGAAATACCCCAAAGC	<i>XbaI</i>
rev-MAP3K7IP2	CAGACTAGTTGATCTGTTGGTGGGTGAA	<i>SpeI</i>
ss-MAP3K7IP2-mut	CCTTAAATAAGAT TCGA AAATAAGAACTATGCACAAAATAGC	
rev-MAP3K7IP2-mut	GTTTCTTATT TCGA ATCTTATTTAAGGCATTGCTTCTCTC	
ss-GPM6B	CAGTCTAGACCTGTTGCATGAAGGAGACA	<i>XbaI</i>
rev-GPM6B	CAGACTAGTCAGCTACAGGTCTGTTTAA	<i>SpeI</i>
ss-GPM6B-mut	ATGTAAAGAT TCGA AAATTTTGCTATGTATAAAATTTTGTAAACC	
rev-GPM6B-mut	CATAGCAAAATTT TCGA ATCCTTACATAAAAAATATTAGAC	
ss-RREB1	CAGTCTAGAAGGTTTCCAGGTCTCGTTT	<i>XbaI</i>
rev-RREB1	CAGACTAGTGAAATTCCTTTGCGCTTCCTG	<i>SpeI</i>
ss-RREB1-mut1	CATTTAA TCGA AAATGTTACCTTTTGATTGGTGTGTGTG	
rev-RREB1-mut1	GGTAACATT TCGA ATTAATGAAAAAGGTCTCTCC	
ss-RREB1-mut2	CCACACGGTGGAT TCGA AAGCCTGTTCCAGGCATTAGCAG	
rev-RREB1-mut2	GGAACAGGCTT TCGA ATCCACCGTGTGGCAAATGCGAG	
ss-c-Myb	CAGGCGGCCGCGACTTCTCCGAGAAGC	<i>NotI</i>
rev-c-Myb	CAGGCGGCCGCAATTAAGTGCAGGTGAAGC	<i>NotI</i>
ss-c-Myb-mut1	CAGGCGGCCGCGACTTCTCCGAGAA TCGA ATGGTG	<i>NotI</i>
ss-c-Myb-mut2	GAAGTAAAAA TCGA ACTACACACGCTAGTCTGTGC	
rev-c-Myb-mut2	GTAGT TCGA ATTTTTTACTTCAGTATTTTACACCAATGC	
ss-PU.1	CAGTCTAGAGCCACGGGACCCTAAGCAA	<i>XbaI</i>
rev-PU.1	CAGACTAGTAAATATGTACATGTTATTAC	<i>SpeI</i>
ss-PU.1-mut1	CAGTGACACTT TCGA AGCTTACAGCTTTCAAACCTAGC	
rev-PU.1-mut1	GCTGTAAGC TCGA AAAGTGTCACTGTGAGTGACACAGC	
ss-PU.1-mut2	GCTTTCAAAC TCGA AAATACAGGCTCTGCCGCAGCTC	
rev-PU.1-mut2	GCCTGTATT TCGA ATAGTTTGAAAGCTGTAAGCTAATGCAG	
ss-PU.1-mutx2	CAGTGACACTTTCGAAGCTTACAGCTTTCAAACCTATTCGAAATACAGGCTCTGCCGCAGCTC	
rev-PU.1-mutx2	GCCTGTATTTCGAATAGTTTGAAAGCTGTAAGCTTCGAAAGTGTCAGTGTGAGTGACACAGC	
ss-UL-28	CAGTCTAGAGGCGAAGTATTGATAGACTC	<i>XbaI</i>
rev-UL-28	CAGACTAGTTCGGCACGACAGATGTGC	<i>SpeI</i>
ss-UL-28-mut	CCAGCCGTCGAATAG TAAGG ATATTTTAGTGTGAAAATG	
rev-UL-28-mut	CATTTTCAACACTAAAATAT CCTTCA CTATTCGACGGCTGG	
ss-UL-32	CTGTCTAGATGCTTGCGGATCA	<i>XbaI</i>
rev-UL-32	CAGACTAGTTGTCGCCGACTCCGTCCAACCTA	<i>SpeI</i>
ss-UL-32-mut	TAAAGCCCCTCT CG CAGTCA CTAGC CGTTTGAAGGTCGCGTATG	
rev-UL-32-mut	ACGCGACCTTCAAA CGGC TA GTG ACT GCG AAGACGGGCTTTATAT	

Tableau 7 : Oligonucléotides utilisés pour le clonage des cibles sauvages et mutantes des microARN.

* Les sites de restriction servant au clonage des amplicons sont soulignés

§ les substitutions pour la construction des cibles mutantes sont indiquées en gras

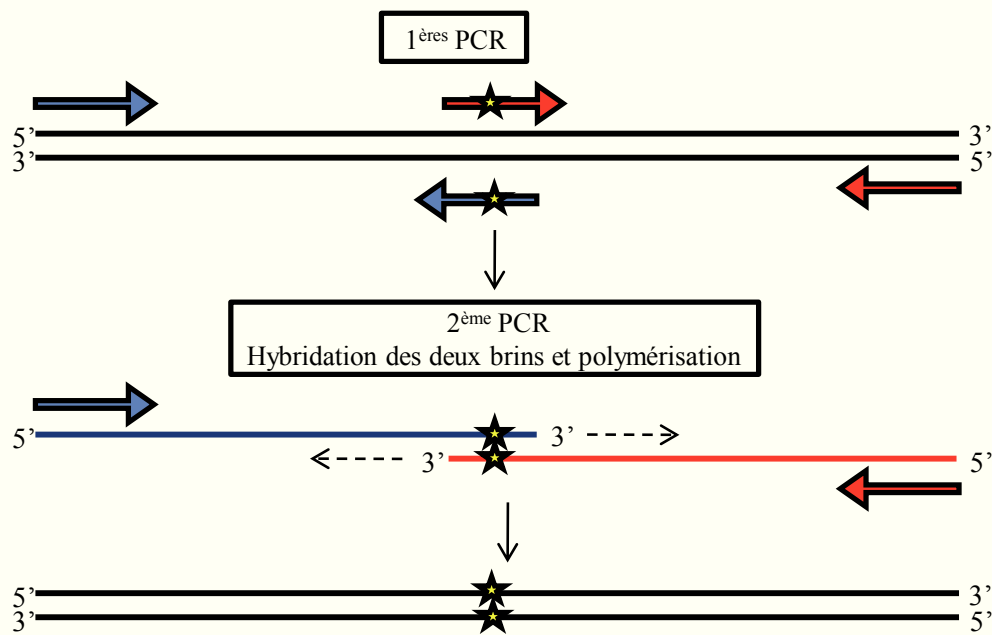


Figure 42 : Elaboration des constructions comportant des mutations internes par mutagenèse dirigée.

Construction	1 ^{ère} PCR		2 ^{ème} PCR
	Matrice	Oligonucéotides	Oligonucléotides
pRL-TK-CEBPβ-mut	ADNg PA-12	ss-CEBPβ/rev-CEBPβ-mut ss-CEBPβ-mut/rev-CEBPβ	ss-CEBPβ/rev-CEBPβ
pRL-TK-MAP3K7IP2-mut	ADNg PA-12	ss-MAP3K7IP2/rev-MAP3K7IP2-mut ss-MAP3K7IP2-mut/rev-MAP3K7IP2	ss-MAP3K7IP2/rev-MAP3K7IP2
pRL-TK-GPM6B-mut	ADNg PA-12	ss-GPM6B/rev-GPM6B-mut ss-GPM6B-mut/rev-GPM6B	ss-GPM6B/rev-GPM6B
pRL-TK-RREB1-M1	ADNg PA-12	ss-RREB1/rev-RREB1-M1 ss-RREB1-M1/rev-RREB1	ss-RREB1/rev-RREB1
pRL-TK-RREB1-M2	ADNg PA-12	ss-RREB1/rev-RREB1-M2 ss-RREB1-M2/rev-RREB1	ss-RREB1/rev-RREB1
pRL-TK-RREB1-M1M2	pRL-TK-RREB1-M1	ss-RREB1/rev-RREB1-M2 ss-RREB1-M2/rev-RREB1	ss-RREB1/rev-RREB1
pRL-TK-c-Myb-M1	ADNg PA-12	ss-c-Myb-M1/rev-c-Myb	
pRL-TK-c-Myb-M2	ADNg PA-12	ss-c-Myb/rev-c-Myb-M2 ss-c-Myb-M2/rev-c-Myb	ss-c-Myb/rev-c-Myb
pRL-TK-c-Myb-M1M2	pRL-TK-c-Myb-M2	ss-c-Myb-M1/rev-c-Myb	
pRL-TK-PU.1-M1	ADNg PA-12	ss-PU.1/rev-PU.1-M1 ss-PU.1-M1/rev-PU.1	ss-PU.1/rev-PU.1
pRL-TK-PU.1-M2	ADNg PA-12	ss-PU.1/rev-PU.1-M2 ss-PU.1-M2/rev-PU.1	ss-PU.1/rev-PU.1
pRL-TK-PU.1-M1M2	ADNg PA-12	ss-PU.1/rev-PU.1-M1M2 ss-PU.1-M1M2/rev-PU.1	ss-PU.1/rev-PU.1
pRL-TK-UL-28-mut	ADNg RB-1B	ss-UL-28/rev-UL-28-mut ss-UL-28-mut/rev-UL-28	ss-UL-28/rev-UL-28
pRL-TK-UL-32-mut	ADNg RB-1B	ss-UL-32/rev-UL-32-mut ss-UL-32-mut/rev-UL-32	ss-UL-32/rev-UL-32

Tableau 8 : Elaboration des constructions comportant des mutations internes par mutagenèse dirigée. Les séquences nucléotidiques des oligonucléotides utilisés sont décrites dans le tableau 7.

enzymes *XbaI* et *SpeI*, ou dans le cas de c-Myb par l'enzyme *NotI*, les amplicons ont été insérés dans le vecteur pRL-TK afin de générer les vecteurs pRL-TK-CEBP β -wt, pRL-TK-GPM6B-wt, pRL-TK-MAP3K7IP2-wt, pRL-TK-RREB1-wt, pRL-TK-c-Myb-wt, pRL-TK-PU.1-wt, pRL-TK-UL-28-wt et pRL-TK-UL-32-wt.

Les mutations des miRE présents au sein des 3'UTR ou des ORF ont été réalisées par mutagenèse dirigée (Figure 42). Deux fragments ont été initialement produits par PCR par l'emploi de deux couples d'amorces spécifiques. Chaque couple est composé d'un oligonucléotide bordant l'insert et d'un oligonucléotide comportant la mutation interne. L'amplification de la séquence mutante complète est ensuite réalisée à l'aide des oligonucléotides terminaux en utilisant comme matrice les deux fragments précédemment amplifiés. Les oligonucléotides et les matrices utilisées pour générer les constructions des 3'UTR possédant les cibles mutées sont présentés dans les tableaux 7 et 8.

Tous les constructions des 3'UTR et ORF mutantes ont été insérés dans le plasmide rapporteur au niveau des sites de restriction *XbaI*/*SpeI* ou *NotI* pour c-Myb, afin de générer les plasmides pRL-TK-RREB1-M1, pRL-TK-RREB1-M2, pRL-TK-RREB1-M1M2, pRL-TK-c-Myb-M1, pRL-TK-c-Myb-M2, pRL-TK-c-Myb-M1M2, pRL-TK-PU.1-M1, pRL-TK-PU.1-M2, pRL-TK-PU.1-M1M2, pRL-TK-UL-28-mut et pRL-TK-UL-32-mut.

4.3. Clonage des protéines UL-28 et UL-32 taguées

La séquence codante des protéines UL-28 et UL-32 a été amplifiée par PCR à partir d'ADNg de la souche GaHV-2 RB-1B au moyen des oligonucléotides « sens » commençant au niveau du codon initiateur des protéines et des oligonucléotides « anti-sens » se terminant juste en amont du codon stop (Tableau 9). Après purification, le produit de PCR a directement été inséré au sein du vecteur pcDNA3.1-V5-His-TOPO (Invitrogen) afin de générer les plasmides pcDNA-UL-28-V5 et pcDNA-UL-32-V5.

4.4. Clonages des constructions promoteurs du cluster1 de microARN

La région promotrice du cluster 1 de microARN de GaHV-2 se situant en position 132013 à 133627 au niveau de l'IR_L de GaHV-2 a été amplifiée par PCR à partir d'ADNg de la souche GaHV-2 RB-1B au moyen des nucléotides M415 et M431 (Tableau 10). Après digestion enzymatique, cette séquence promotrice a été insérée en orientation sens au niveau des sites *KpnI*/*HindIII* du plasmide rapporteur pGL3basic afin de générer le plasmide pGL3-p-miR-M9-M4.

Nom	Séquence 5' → 3'	Position [§]
ss-UL-28-ORF	<u>ATG</u> *TTGGGAATGTCTCATAACCGGTT	66108-66083
rev-UL-28-ORF	GATGGGGGCGTGGCTGTTGGTAATCT	63730-63735
ss-UL-32-ORF	<u>ATGGCCA</u> ACCGCCCTACAGAGTTGGC	76574-76559
rev-UL-32-ORF	CACGTAGACTCCTAATGTATGCTCGA	74652-74677

Tableau 9 : Oligonucléotides utilisés pour le clonage des protéine taguées UL-28-V5 et UL-32-V5.

* les codons d'initiations des protéines UL-28 et UL-323 sont soulignés

§ la position est donné en fonction du génome GaHV-2 Md5

Nom	Séquence 5' → 3'	Orientation [#]	Site de restriction	Position [§]
M415	GGTACC* <u>AACGATTGCGGAAGTACGGC</u>	sens	<i>KpnI</i>	132013-132032
M442	<u>GGTACCTGTTCTGTGCGCTTCTCAAC</u>	sens	<i>KpnI</i>	132307-132326
M443	<u>GGTACCAAGTCGCTGTTCCCATGAGA</u>	sens	<i>KpnI</i>	132626-132645
A78	<u>GGTACCGTGCTGATTATCATCTCATTAC</u>	sens	<i>KpnI</i>	132768-132790
M444	<u>GGTACCTATGTCGAGAGGCCCTGTAA</u>	sens	<i>KpnI</i>	132886-132905
M445	<u>GGTACCCCGGAGTTCACTGTATCGTA</u>	sens	<i>KpnI</i>	133388-134407
M431	<u>AAGCTTGTC</u> AATGTGATCGCATACGG	anti-sens	<i>HindIII</i>	133626-133607
M542	<u>AAGCTTTACGATA</u> CAGTGAACCTCCGG	anti-sens	<i>HindIII</i>	133407-133388
M476	<u>AAGCTTAAGCTAAAAT</u> CCGGGCATC	anti-sens	<i>HindIII</i>	132699-132680

Tableau 10 : Oligonucléotides utilisés pour le clonage des construction promotrices de la région p-miR-M9-M4.

* les sites de restriction sont soulignés

[#] l'orientation est donné en fonction du brin codant les miARN de l'IR_L de GaHV-2

§ la position est donné en fonction du génome GaHV-2 Md5

Neuf promoteurs tronqués de p-miR-M9-M4 ont été obtenus par PCR à partir du plasmide pGL3-p-miR-M9-M4. Les cinq promoteurs tronqués dans la partie 5' de p-miR-M9-M4 nommés Trq-1, Trq-2, Trq-3, Trq-4 et Trq-5 ont été respectivement amplifiés avec les amorces « sens » M442, M443, A78, M444 et M445 et de l'amorce « anti-sens » M431 (Tableau 10). Les deux promoteurs tronqués dans la partie 3' de p-miR-M9-M4 nommés p-court1 et p-court2 ont été respectivement amplifiés avec l'amorce « sens » M415 et des amorces « anti-sens » M542 et M476 (Tableau 10). Pour finir, les deux promoteurs tronqués dans les parties 5' et 3' Trq-3-court et Trq-4 court ont été respectivement amplifiés à l'aide des amorces « sens » A78 et M444 et de l'amorce « anti-sens » M542.

L'ensemble de ces séquences promotrices ont été digérés par les enzymes *KpnI* et *HindIII* et ont été insérées dans le vecteur pGL3-basic afin de générer les vecteurs pGL3-Trq-1, pGL3-Trq-2, pGL3-Trq-3, pGL3-Trq-4, pGL3-Trq-5, pGL3-p-court-1, pGL3-p-court-1, pGL3-Trq-3-court, pGL3-p-Trq-4-court, et pGL3-Trq-4.5-court.

5. Expression des plasmides en système cellulaire eucaryote

Le kit Dual-Luciferase Reporter Assay System (Promega) a été utilisé afin de tester l'efficacité transcriptionnelle des différents promoteurs ainsi que l'effet des miARN sur les cibles potentielles. Dans le cas de l'étude de l'efficacité transcriptionnelle des différents promoteurs, le plasmide pRL-TK codant pour la luciférase *Renilla* sous le promoteur de la thymidine kinase de HHV-1 est cotransfecté afin de standardiser l'efficacité de transfection. Le plasmide pcDNA-MLuc, codant la luciférase *Firefly* en aval du promoteur CMV, et le plasmide pGL3-Basic ont été transfectés lors de chaque essai afin de servir respectivement de témoins positifs et négatifs. A l'inverse, lors de l'étude de l'effet des miARN sur les cibles cellulaires et virales, la standardisation de l'efficacité de transfection est réalisée en rapportant l'activité luciférase *Renilla* à l'activité luciférase *Firefly* correspondant au plasmide pcDNA-MLuc cotransfecté dans chaque essai.

5.1. Lipofection des cellules adhérentes

Vingt-quatre heures avant transfection, les cellules DF-1 sont mises en culture en plaque de 24 puits à une densité de 6.10^4 cellules par puits. La lipofection étant optimale en milieu dépourvu de sérum, les tapis cellulaires sont lavés 2 fois en milieu OptiMEM (Invitrogen), puis recouverts par 200 μ l d'OptiMEM. Les plasmides (1,2 μ g des vecteurs d'expression des miARN, 60 ng des différentes constructions contenant les cibles sauvages ou mutantes et 15 ng de plasmide pcDNA-MLuc) sont

dilués dans 100 µl de milieu OptiMEM. En parallèle, 4 µl d'agent de transfection Lipofectamine 2000 (Invitrogen) sont dilués également dans 100 µl d'OptiMEM. Après 20 min d'incubation à température ambiante, les deux solutions sont mélangées et incubées 20 min à température ambiante. Cette suspension de liposomes associés aux plasmides est alors déposée sur les cellules. Après 6 heures d'incubation à 41°C, la suspension est remplacée par 1 ml de milieu de culture DMEM (Lonza) additionné de 10% de SVF et 5% de SP. Les cellules sont alors incubées à 41°C durant 24 heures avant la mesure de l'activité luciférase. Il faut noter que lors de chaque essai, chaque combinaison de plasmide est systématiquement transfectée dans 3 puits distincts de P24.

5.2. Electroporation des cellules en suspension

5.2.1. Electroporateur EasyjecT Plus

Les cellules PA.9, MSB-1 et 54-O sont transfectées à l'aide de l'électroporateur EasyjecT Plus (Equibio). Après 2 lavages en milieu sans sérum, 8 millions de cellules sont diluées dans 800 µl de milieu RPMI 1640 puis placées dans une cuvette d'électroporation de 4 mm à électrodes en aluminium (Eurogentec). Quarante microgrammes de plasmide rapporteur et 500 ng de plasmide contrôle (pRL-TK : PA-9 et pcDNA-MLuc : MSB-1 et 54-O) sont ajoutés aux cellules qui sont alors soumises à une impulsion électrique simple de 400V et 1500 µF. Après électroporation, les cellules sont mises en culture en plaque 6 puits contenant 2,5 ml de milieu RPMI 1640 additionné de 10% de SVF et 5% de SP, puis incubées à 41°C.

5.2.2. Electroporateur Amaxa

Les cellules MSB-1 et DT-40 sont électroporées à l'aide de l'électroporateur Amaxa (Amata Biosystems). Dans le cas de l'étude de l'efficacité transcriptionnelle des promoteurs, deux millions de cellules MSB-1 sont mélangées avec 100 µL de la solution de nucléofection T contenant 1µg de plasmide rapporteur et 50ng de plasmide contrôle pRT-TK. Lors de l'étude de l'effet des miARN sur les cibles cellulaires et virales, deux millions de cellules DT-40 sont mélangées avec 100 µL de la solution de nucléofection T contenant 4µg de plasmide d'expression des miARN, 200ng du vecteur rapporteur et 50ng de plasmide contrôle pcDNA-MLuc. Après électroporation les cellules MSB-1 et DT-40 sont immédiatement reprises dans 500µL de milieu RPMI1640 puis déposées dans un puits d'une plaque de type P12 contenant 2mL de milieu RPMI1640 additionné de 10% de SVF et 5% de SP, puis incubées à 41°C.

5.3. Mesure de l'activité luciférase

Vingt-quatre heures après transfection les tapis cellulaires sont lavés 2 fois en PBS 1X. Les cellules DF-1 sont lysées par dépôt de 100 µl de tampon de lyse (Passive Lysis Buffer 1X) directement sur le tapis cellulaire alors que les cellules MSB-1, 54-O et DT-40 sont lysées après 5 min de centrifugation à 300g, par reprise du culot cellulaire dans 50 µl de ce même tampon de lyse. Les lysats cellulaires sont alors incubés 15 min sous agitation douce à température ambiante. Après clarification du lysat par 30 secondes de centrifugation à 16000g, 10 µl de surnageant est transféré dans une plaque P96 pour « luminométrie » (Berthold Technologies).

L'activité des luciférases *Firefly* et *Renilla* contenues dans chaque extrait cellulaire est alors mesurée à l'aide d'un luminomètre Mithras LB940 (Berthold Technologies), selon le principe du kit Dual luciferase (Promega). Ainsi, après injection de 50 µl du substrat de la luciférase *Firefly* (LAR II), l'automate mesure l'activité lumineuse associée au vecteur pGL3 ou pcDNA-MLuc puis injecte 50 µl du substrat de la luciférase *Renilla* (Stop & Glo) bloquant la luciférase *Firefly* et activant la luciférase *Renilla* associée au vecteur pRL-TK.

6. Analyse des ARN totaux

6.1. Extraction des ARN totaux

Les ARN totaux sont extraits à l'aide du réactif TRIzol selon les conditions décrites par le fournisseur (Invitrogen). Cinq millions de cellules sont lavées 2 fois par 1 ml de PBS 1X, puis lysées par 1 ml de TRIzol. Après incubation pendant 5 min à 30°C, 200µl de chloroforme est additionné, puis les extraits sont « vortexés » et incubés 2-3 min à 30°C. Après 15 min de centrifugation à 12000g à 4°C, la phase supérieure contenant les ARN est prélevée. Les ARN sont alors précipités par 500µl d'isopropanol. Après 10 min d'incubation à 30°C et une centrifugation de 10 min à 12000g à 4°C, le culot d'ARN est lavé avec 1 ml d'éthanol 75 % et séché à température ambiante. Il est ensuite repris dans 30 µl d'eau et incubé 5 min à 60°C afin d'optimiser la remise en suspension. Les ARN ainsi extraits sont soit conservés à -80°C, soit directement traités à la DNase.

Afin d'éliminer toute trace d'ADN génomique résiduelle, un traitement à la DNase est nécessaire. La réaction est réalisée dans un volume final de 100 µl comprenant 40 unités de RNAsin (Promega), 4 unités de RQ1 DNase RNase-Free (Promega), 40 mM de Tris-HCl pH 8, 6 mM de MgCl₂, 10 mM de MgSO₄, 10 mM de CaCl₂ et 85 µl d'ARN totaux. Après une incubation d'une heure

Nom	Orientation*	Séquence 5' → 3'	position§
A13	sens	CATACAGGATCTGGCTGCGCGCCAG	133055-133079
A14	sens	GTAGCGGGGACGTTTCATTGTCTTTG	132057-132081

Tableau 11 : Oligonucléotides utilisés pour l'amplification du brin inverse aux miARN lors de la reverse transcription.

* l'orientation est donné en fonction du brin codant les miARN de l'IR_L de GaHV-2

§ la position est donné en fonction du génome GaHV-2 Md5

Nom	Orientation [#]	Séquence 5' → 3'	position [‡]
M854	sens	ATGCTCTCTCAGAATGGCACGGC	128473-128495
M835	sens	GTCTACTTCCACAAGGGGGTG	129194-129214
M855	sens	CGGCAGAAAGAAACCCCTTCTCAG	129812-129834
A230	sens	GACTGCGGACCCACTTCCATC	130915-130936
M830	sens	GAAAGTTTGCCGTCCGCCTG	131616-131635
M774	sens	AACGATTGCGGAAGTACGGC	132013-132032
M852	sens	GCCCTGTTCGGCCAATCGAATTC	133017-133040
M849	sens	CTTTTCGCCCCTGATCCCTCTACCAG	133117-133140
M445	sens	GGTACCCCGAGTTCACTGTATCGTA	133388-133407
M421 [§]	sens	GTCTTCGTGACACCGCGTAT	134158-134177
A50 [§]	sens	CGTGAAAGTCACTAATCGGC	138347-138366
M829	anti-sens	GACCGATTTACAGGATGAGGCTG	132138-132116
M479	anti-sens	CAGGGCCTCTCGACATAAAG	132902-132883
A72	anti-sens	CGCACTTTGGTTCTAAATGTC	133177-133157
M478	anti-sens	GTATGTGAGTCGGGGCACTT	133452-133433
M416*	anti-sens	GTCAATGTGATCGCATACGG	133627-133608
A64	anti-sens	CATTTTCCCATCCTGCTCAC	133839-133820
A65	anti-sens	GGCCACAGTAAGTCTGGCTC	134015-133996
M853	anti-sens	ATACGCGGTGTCACGAAGACCGA	134177-134155
A66	anti-sens	ATTCAAACAAGCCCAGAGC	134353-134334
M790	anti-sens	CGATGGTTTCCAGCGATCA	134458-134439
A376	anti-sens	GGAGCAATTAGTTTATTTAAGAGG	134655-134633
M419	anti-sens	CTAGGAATCAGTGTGCGGAA	134806-134787
A15*	anti-sens	GTGCCCCTTCTCCCTGGTATAC	134849-134826
M420*	anti-sens	CTCCTCAGATAGGCCGTCAG	135031-135012
770	anti-sens	AACCGGAGCAATGTGGAGCGTTAGG	135368-135344
657	anti-sens	GGGTACACGGCTCGGTAACAG	135665-135645
636	anti-sens	TGCTGCTGAGAGTCACAATGC	135913-135893
M825	anti-sens	TGCCACATCGTAGAGAAAGC	137028-137009
M826	anti-sens	CGCAGATACACTCCACCGTAC	137489-137469
M827	anti-sens	GACCCCTGCATAGCAGAAAAC	137764-137744
M04*	anti-sens	TCAGTCTCTGTCAGTGTATACCCGG	138433-138408

Tableau 12 : Oligonucléotides utilisés pour l'amplification des banques d'ADNc de la région IRL de GaHV-2.

* Oligonucléotides utilisés lors des 5'RACE PCR

§ oligonucléotides utilisés lors des 3'RACE PCR

l'orientation est donné en fonction du brin codant les miARN de l'IR_L de GaHV-2

‡ la position est donné en fonction du génome GaHV-2 Md5.

à 37°C, la DNase est inactivée par une incubation de 10 min à 65°C. Après 10 min dans la glace, l'ARN est extrait par un volume de phénol-chloroforme-alcool isoamylique (50-49-1) puis précipité par 3 volumes d'alcool à 100% en présence de 0,3 M d'acétate de sodium. Après centrifugation le culot d'ARN est soit conservés à -80°C en alcool à 80%, soit repris en eau RNase free avant d'être transcrit en cDNA.

6.2. Transcriptase inverse

Les ADNc sont synthétisés à partir d'un microgramme d'ARN (estimé par mesure de la D.O à 260 nm) dilué dans un volume final de 20µl. Les ARN totaux sont dénaturés par chauffage 10 min à 50°C en présence d'un mélange de 50 pmol d'oligonucléotide dT et de 50 pmol d'hexamères aléatoires ou d'amorces spécifiques (Tableau 11), puis placés 5 min dans la glace. Le mélange réactionnel composé de 400 unités de la reverse transcriptase superscript III (Invitrogen), 100 µM de chaque désoxy-ribonucléotide, 50mM de Tris-HCl pH 8,3, 75 mM de KCl, 3 mM de MgCl₂, 10 mM de DTT, est ensuite ajouté et les échantillons sont incubés 1 heure à 37°C. La réaction est arrêtée par 5 min d'incubation à 85°C. Les ADNc ainsi obtenus sont conservés à -20°C.

6.3. Amplification génique sur ADNc

Les amplifications géniques sur ADNc sont effectuées au moyen d'oligonucléotides spécifiques (Tableau 12, 13). Les mélanges réactionnels spécifiques de chaque gène d'intérêt, d'un volume final de 50 µl, sont soumis à 25 cycles d'amplification (1 minute, 94°C; 1 minute, 55°C; 1 minute, 72°C) suivis d'une étape de 10 min à 72°C permettant de terminer la synthèse des brins.

La PCR nichée a été réalisé en utilisant 1/50^{ème} du produit de PCR précédent comme matrice des différentes amplifications, au moyen d'un couple d'oligonucléotides spécifiques. Les mélanges réactionnels spécifiques de chaque gène d'intérêt, d'un volume final de 50 µl, sont soumis à 25 cycles d'amplification (1 minute, 94°C; 1 minute, 55°C; 1 minute, 72°C) suivis d'une étape de 10 min à 72°C permettant de terminer la synthèse des brins.

6.4. RACE PCR

Les RACE-PCR (Rapid Amplification of cDNA Ends) 3' et 5' sont réalisées à l'aide du kit GeneRacer(Invitrogen).

Nom	Orientation*	Séquence 5' → 3'	position #
A5	sens édité	GCCTTTCGATCGTG GT AT CC §	133144-133163
A6	anti-sens édité	AGTGAGAG TT GT CTCTCGTG CG	133346-133327
A7	anti-sens édité	TGCGGAGATGT GT GT G GGT G	133317-133298
A8	sens édité	GCGCGGG CA AGG CAAT CACC	133173-133192
A9	sens non édité	GCTTTTCGATCGTGACATT	133144-133163
A10	anti-sens non édité	AATGAGAATTATCTCTCGTGCA	133346-133327
A11	anti-sens non édité	TACGGAATGTATATGAGTA	133317-133298
A12	sens non édité	GTGCGGGTAAGGTAATCATT	133173-133192

Tableau 13 : Séquence et position des oligonucléotides utilisés pour lors de la recherche des transcrits édités.

* l'orientation est donné en fonction du brin codant les miARN de l'IR_L de GaHV-2

§ les nucléotides en rouge correspondent aux bases éditées

la position est donné en fonction du génome GaHV-2 Md5

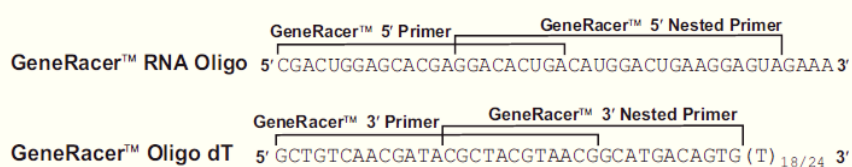


Figure 43 : Séquence des oligonucléotides ARN utilisés lors des RACE-PCR.

Les oligonucléotide GeneRacer RNA Oligo et GeneRacer Oligo dT sont utilisés respectivement lors des 5' et 3' RACE-PCR.

Pour la 3' RACE, aucun traitement préalable à la synthèse des ADNc n'est nécessaire. Par contre, en vue de la 5' RACE, 5 µg d'ARN total sont tout d'abord traités par la CIP (Calf Intestinal Phosphatase) afin de déphosphoryler les ARN non cappés (ARNm tronqués et les non messagers). Les ARNm ne sont pas affectés car les phosphates sont protégés par la coiffe en 5'. Les ARN sont ensuite extraits par addition d'un volume de phénol-chloroforme-alcool isoamylique (50-49-1). Après homogénéisation et centrifugation 5 min à 13000g, la phase aqueuse est récupérée, additionnée d'acétate de sodium (0,3M final), de 20µg de mûsse glycogène et 2,5 volumes d'éthanol 95% puis incubée dans de la carboglace pendant 10 min. Après 20 min de centrifugation à 13000g à 4°C le culot d'ARN est lavé à l'éthanol 75 % à nouveau centrifugé avant d'être repris dans de l'eau. Il faut noter que cette étape d'extraction et de purification de l'ARN est réalisée systématiquement entre chaque étape du protocole de RACE PCR (Invitrogen).

La coiffe est ensuite enlevée de l'extrémité 5' des ARNm par un traitement à la TAP (Tobacco Acid Pyrophosphatase), qui coupe entre le phosphate alpha et bêta, maintenant ainsi le phosphate alpha en 5' nécessaire pour l'étape de ligation suivante. L'oligonucléotide ARN 5' GeneRacer (figure 43) est alors ligué sur les ARNm « dépourvus de coiffe » grâce à l'action d'une T4 RNA ligase, afin de positionner sur l'extrémité 5' une séquence connue en amont de la région d'intérêt utilisable dans les PCR ultérieures.

Les ADNc en vue de la 5' ou de la 3' RACE PCR sont ensuite synthétisés selon le même protocole, l'unique différence résidant dans les amorces utilisées pour initier leur synthèse : un mélange d'amorces hexanucléotidiques (50ng) et d'oligonucléotide dT (450ng) pour la 5'RACE ou l'oligonucléotide dT (900ng) du kit GeneRacer pour la 3'RACE. Les structures secondaires des ARN sont dénaturées par incubation à 65°C pendant 5 min puis 2 min dans la glace. Dans un premier temps, les amorces ainsi que le mélange de 200 µM chaque désoxy-ribonucléotide sont ajoutés à l'ARNm. Après complétion du reste du mélange réactionnel de composition finale : 5mM de dithiothreitol (DTT), de 40 unités de RNasin (Promega), de 200 unités de SuperScript III Reverse Transcriptase (Invitrogen) et de son tampon spécifique (250mM tris-HCl, 375 mM KCl, 15mM MgCl₂), le mélange réactionnel est incubé une heure à 52°C. La réaction est stoppée par 15 min d'incubation à 70°C. La matrice d'ARN est ensuite dégradée par l'action de la RNase H pendant 20 min à 37°C. Les ADNc obtenus sont alors stockés à -20°C ou directement utilisés pour une réaction de PCR.

L'amplification est effectuée sur les cDNA préparés pour la 3' et la 5' RACE-PCR à l'aide du kit GeneRacer (Invitrogen) avec des couples d'amorces composés d'une amorce spécifique (Tableau 12) et d'une amorce GeneRacer 5' ou 3' (Tableau 14). Les mélanges réactionnels sont élaborés dans un volume final de 50µl contenant 1,25 unité de GoTaq Flexi DNA polymérase (Promega), 150µM de chaque désoxy-ribonucléotide, 10 mM de Tris-Hcl pH 9, 50mM de KCl, 0,5 mM de MgCl₂, 0,1 % de Triton X-100, 400 µM d'amorce GeneRacer 5' ou 3', 300 µM d'amorces spécifiques et 1 µl d'ADNc matriciel. Une dénaturation de 3 min à 94°C suivie de 25 cycles d'amplification sont ensuite effectués

Nom	Séquence 5' → 3'	Tm
GeneRacer 5'Primer	CGACTGGAGCACGAGGACACTGA	74°C
GeneRacer 5'Nested Primer	GGACACTGACATGGACTGAAGGAGTA	78°C
GeneRacer 3'Primer	GCTGTCAACGATACGCTACGTAACG	76°C
GeneRacer 3'Nested Primer	CGCTACGTAACGGCATGACAGTG	72°C

Tableau 14 : Amorces du kit GeneRacer (Invitrogen) utilisées lors des RACE-PCR.

Nom	Séquence 5' → 3'
as-mdv1-miR-M4-5P	ACGAAGGGTTCCGATACAGCATTAA
as-mdv1-miR-M4-5P	GGTCATGCTGTCAGAACCATT
as-gga-miR-155	CCCCTATCACGATTAGCATTAA
as-gga-miR-21	TCAACATCAGTCTGATAAGCTA

Tableau 15 : Séquence des oligonucléotides utilisés pour la détection des miARN par Northern blot.

(30sec : 94°C / 30sec : 62°C / 2min : 72°C). Une dernière étape d'élongation de 7 min à 72°C termine la réaction de PCR. Des PCR pseudo-nichées, avec la même amorce qu'en 1^{ère} PCR et l'amorce 5' nichée du kit GeneRacer (Tableau 14), sont effectuées à partir de 1/50^{ème} de la première PCR. Après une dénaturation à 94°C pendant 3 min, 25 cycles sont réalisés (30sec : 94°C / 30sec : 62°C / 2min : 72°C) avant une dernière élongation de 7 min à 72°C.

6.5. Northern -blot

Après extraction, 15µg d'ARN total ont été déposés par puits d'un gel d'acrylamide/bisacrylamide (19:1) contenant 7M d'urée dans un tampon TBE. Après migration de 1 heure à 50mA le gel est plongé dans un bain contenant du bromure d'éthidium afin de vérifier la présence des ARN, puis les ARN sont transférés par électrotransfert sur une membrane Nylon Plus (Amersham) pendant 90 min à 350mA en présence de TBE 0,5X. Après séchage les ARN sont fixés à la membrane par exposition aux UV dans un « crosslinker » (Stratalinker, Stratagene). Parallèlement, les sondes sont marquées à l'aide de la T4 PNK (T4 Polynucléotide Kinase ; Promega) au sein d'un mélange réactionnel de 20µL contenant 10 unité de PNK, 20mM d'oligonucléotide à marqué, en présence de 25µCi de γ ATP puis incubé pendant 30min à 37°C. La réaction est ensuite stoppée par ajout de 30µL d'EDTA à 30mM à pH 8. La membrane est ensuite incubée dans un tampon de préhybridation pendant une heure à 50°C puis dans du tampon Perfect Hyb TM Plus (Sigma) pendant 12 heures à 50°C en présence de 20nM oligonucléotides ADN complémentaires des miARN et marqués à leur extrémité 5' par un ^{32}P (Table 15). La membrane est alors rincée à l'aide de 10 ml de tampon wash1 (SSC 5X, 0,1% SDS), puis deux lavages de 10 min sont effectués à l'aide du tampon wash1, suivi de deux lavages de 10 min avec le tampon wash2 (SSC 1X, 0,1% SDS permettant d'éliminer les hybridations non parfaitement complémentaires. Les membranes sont ensuite révélées à l'aide d'un phosphorimager Storm 840.

7. Western-Blot

Deux jours après les transfections, les cellules sont lysées à l'aide de 20µL de tampon de lyse (15mM Tris, 15mM NaCl, 20mM EDTA, 1%SDS, 5% beta-mercaptoethanol, 20% glycerol et de bleu de bromophénol). Après avoir été portées à ébullition, les protéines sont séparées par migration électrophorétique en gel de polyacrylamide dénaturant (SDS-PAGE : Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis). Le gel de séparation à 7,5% de polyacrylamide est composé de

1,25ml de tampon Tris 4X (tampon Lower : Tris 181,7 g/L, SDS 4 g/L, pH 8,8), 1 ml d'acrylamide:bisacrylamide 37,5:1 à 40% (BioRad), 15 µl d'ammonium persulfate à 10%, 5 µl de TEMED (N, N, N', N'-Tetramethylethylenediamine ; Sigma) et 2,75 ml d'eau. Le gel de concentration est composé de 500 µl de tampon Tris 4X (tampon Upper : Tris 60,6 g/L, SDS 4 g/L, pH 6,8), 270 µl d'acrylamide:bisacrylamide 37,5:1 à 40% (BioRad), 10 µl d'ammonium persulfate à 10%, 2,5 µl de TEMED et 1,25 ml d'eau. Les extraits protéiques, préalablement portés à ébullition pendant 5 min avec du tampon de dissociation 5X (25% β-mercaptoéthanol, 50% glycérol, SDS 100 g/l, 25% tampon Upper 4X, bleu de bromophénol), sont soumis à migration dans du tampon Tris-Glycine 1X (Tris 3,027 g/l, Glycine 14,41 g/l) à 80 V dans le gel de concentration puis à 200 V dans le gel de séparation. Après séparation des protéines par SDS-PAGE le gel est immergé 45 min dans une solution d'urée 4,5 M. Les protéines sont alors transférées sur membrane de nitrocellulose par transfert passif dans un tampon de transfert 1X (Tris 0,01 M-pH 7, NaCl 0,05 M, EDTA 2 mM) pendant une nuit. La membrane est ensuite saturée pendant au moins 3h en tampon Tris-NaCl Tween (TNT ; Tris 20 mM-pH 7,4, NaCl 140 mM, Tween 0,05%) contenant 3% de lait écrémé (Régilait). La première incubation est réalisée avec l'anticorps primaire dilué au 1/5000^{ème} dans du tampon TNT contenant 3% de lait écrémé durant 1 h sous agitation douce à température ambiante. La membrane est ensuite lavée 3 fois avec du tampon TNT contenant 1% de lait écrémé. L'incubation avec l'anticorps secondaire couplé à la peroxydase (Pierce) dilué au 1/2500^{ème} dans du tampon TNT contenant 3% de lait écrémé est effectuée durant une heure sous agitation douce à température ambiante. La membrane est alors lavée 3 fois avec du tampon TNT contenant 1% de lait écrémé puis 2 fois avec du tampon TNT dépourvu de lait. Les protéines sont alors révélées par ajout du substrat de la peroxydase (Super Signal West Femto Maximum Sensitivity Substrate, Pierce) et capture de l'image à l'aide d'un lecteur de chimioluminescence (Vilber Lourmat). La masse moléculaire approximative des protéines est déterminée à l'aide du marqueur de taille Precision Plus Protein Standard Dual Color (250, 150, 100, 75, 50, 37, 25, 20 kDa, BioRad).

RÉSULTATS

1. Etude fonctionnelle de mdv1-miR-M4.

1.1. Comparaison des séquences de mdv1-miR-M4 et gga-miR-155.

L'alignement des pré-miARN de gga-miR-155 et de mdv1-miR-M4 fait apparaître un faible taux de similitude globale de l'ordre de 50% entre les 2 pré-miARN (Figure 44). Deux régions sont particulièrement conservées au sein des miARN-5P, l'intégralité des séquences « seed » et une courte région de six nucléotides dans la partie 3' des miARN matures gga-miR-155-5P et mdv1-miR-M4-5P (nucléotides 10-13 et 15-16 chez gga-miR-155 et nucléotides 11-16 chez mdv1-miR-M4-5P). L'alignement des régions 3' des deux pré-miARN fait ressortir une région conservée entre la boucle du pré-miARN cellulaire et une partie de la séquence « seed » de son orthologue viral. Il faut remarquer que les séquences « seed » sont différentes entre les deux miARN-3P. Cependant, comme nous avons localisé le brin passager de gga-pré-miR-155 (gga-miR-155-3P) uniquement par analogie avec les modèles humains et murins, sa localisation précise demeure incertaine. En effet, il n'a jamais été identifié dans les différentes banques de miARN chez le poulet, notamment au sein des six banques plus exhaustives récemment réalisées au laboratoire par séquençage Solexa. De même, le brin passager de gga-pré-miR-155 n'a été identifié qu'à un taux très faible dans les modèles humain et murin (environ 180 fois moins présents que le miARN-5P). A ce jour, il semblerait donc que la différence majeure entre le gga-pré-miR-155 et le mdv1-pré-miR-M4 résiderait dans le fait que le miARN-3P de l'orthologue viral serait exprimé contrairement à son équivalent cellulaire.

1.2. Comparaison du profil d'expression de mdv1-miR-M4 et gga-miR-155.

Plusieurs approches ont été conduites au sein du laboratoire pour évaluer le taux d'expression des miARN au cours de la lymphomagenèse induite par GaHV-2. Dans un premier temps, des banques ont été réalisées suite au clonage des petits ARN extraits de PBL de poulet infectés par la souche hyper-virulente GaHV-2 RB-1B. Entre autres, ces banques de miARN ont révélé un fort taux d'expression du mdv1-miR-M4-5P dans les PBL des poulets infectés, le classant comme le premier ou deuxième miARN viral le plus exprimé. Plus précisément, à j21 et à j31 p.i il représentait respectivement 6,3% et 7,6% des miARN viraux, ces miARN viraux représentant respectivement 10% et 15% de l'ensemble des miARN séquencés. Parallèlement, le mdv1-miR-M4-3P représentait respectivement 0,6% et 0,4% des miARN viraux et gga-miR-155 moins de 0,1% des miARN

cellulaires extraits. Par ailleurs, les banques de miARN établies à partir des petits ARN extraits des lignées lymphoblastoïdes MSB-1 et 54-O ont confirmé les observations précédentes réalisées pour les temps tardifs de l'infection des poulets, correspondant à la lymphomagenèse. En effet, dans ces lignées cellulaires établies à partir de tumeurs induites par GaHV-2, le miARN mdv1-miR-M4-5P représentait 15% et mdv1-miR-M4-3P 1 à 2% des miARN viraux (représentant 40% des miARN totaux) alors que le miARN gga-miR-155 n'était pas détecté. Enfin, il faut noter que ce faible taux d'expression de gga-miR-155 a été confirmé suite au séquençage exhaustif de six banques de miARN correspondant à des poulets non infectés ou au 29^{ème} jour d'infection par la souche vaccinale non-oncogène GaHV-2-CVI988 ou par la souche hyper-virulente GaHV-2-RB-1B, à l'aide de la technique Solexa. En effet, par cette technique qui a permis le séquençage de 5 à 15 millions de miARN par banque, le taux de gga-miR-155 a été estimé à 0,4% des miARN séquencés.

Afin de valider les résultats obtenus dans ces banques, le taux d'expression des trois miARN a été évalué par Northern Blot à partir d'ARN totaux extraits de différentes cellules aviaires : a) non infectées par GaHV-2 (DT40 et PBL non infectés), b) PBL prélevés à J31 sur des poulets infectés par la souche GaHV-2 RB-1B et c) deux lignées lymphoblastoïdes MSB-1 et 54-O (Figure 45). Le miARN gga-miR-21, que nous savions fortement exprimé dans les cellules lymphoïdes infectées et non infectées suite à l'analyse des différentes banques, a été retrouvé dans l'ensemble des extraits d'ARN, tandis que les miARN mdv1-miR-M4 et gga-miR-155 présentaient des profils différents (Figure 45). En effet, mdv1-miR-M4-5P est retrouvé dans les PBL de poulets infectés par la souche GaHV-2 RB-1B et dans les deux lignées tumorales, tandis que gga-miR-155 est exclusivement détecté dans les PBL non infectés. Enfin, le miARN mdv1-miR-M4-3P, ne semble détecté que dans les lignées cellulaires MSB-1 et 54-O, présentant un fort taux d'expression des miARN viraux.

1.3. Recherche de gènes cellulaires ciblés par mdv1-miR-M4-5P et gga-miR-155.

La recherche de cibles potentielles dans les 3'UTR des gènes de poulet, effectuée à l'aide du logiciel TargetScan custom 4.0 (<http://www.targetscan.org/>) (Lewis *et al.*, 2005) a fait ressortir 75 cibles potentielles pour les séquences « seed » des miARN mdv1-miR-M4-5P et gga-miR-155 (Figure 46). Parmi les 75 cibles prédites, six ont été sélectionnées en fonction de leurs implications dans la régulation de la différenciation des lymphocytes, la réponse immunitaire et l'oncogenèse (Figure 46 et 47). Ainsi, six 3'UTR de gènes ont été sélectionnés, trois possédant un seul site de fixation potentiel (miRE : miARN responsive element), (*MAP3K7IP2* (mitogen-activated protein kinase kinase kinase 7 interacting protein 2), *CEBPβ* (CCAAT/enhancer binding protein beta) et *GPM6B* (glycoprotein M6B)), et trois possédant deux miRE (*PU.1* (Ets-family member PU.1), *c-Myb* (v-myb myeloblastosis

Orthologue du gène chez l'humain	Nom du gène	Sites conservés				Sites peu conservés			
		total	8mer	7mer-m8	7mer-1A	total	8mer	7mer-m8	7mer-1A
ZNF652	zinc finger protein 652	2	1	1	0	1	0	1	0
TSHZ3	teashirt family zinc finger 3	3	1	0	2	0	0	0	0
ARID2	AT rich interactive domain 2 (ARID, RFX-like)	2	1	1	0	0	0	0	0
RREB1	ras responsive element binding protein 1	1	1	0	0	1	1	0	0
TLE4	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)	1	1	0	0	0	0	0	0
ZIC3	Zic family member 3 heterotaxy 1 (odd-paired homolog, Drosophila)	1	1	0	0	0	0	0	0
DYNC1H1	dynein, cytoplasmic 1, intermediate chain 1	1	1	0	0	0	0	0	0
GPM6B	glycoprotein M6B	1	1	0	0	0	0	0	0
CEBPB	CCAAT/enhancer binding protein (C/EBP), beta	1	1	0	0	0	0	0	0
ACTA1	actin, alpha 1, skeletal muscle	1	1	0	0	0	0	0	0
FBXO33	F-box protein 33	1	1	0	0	0	0	0	0
TRPS1	trichorhinophalangeal syndrome I	1	0	0	1	2	1	1	0
RNF123	ring finger protein 123	1	1	0	0	0	0	0	0
BACH1	BTB and CNC homology 1, basic leucine zipper transcription factor 1	1	1	0	0	0	0	0	0
LRP1B	low density lipoprotein-related protein 1B (deleted in tumors)	1	0	1	0	1	0	1	0
RCOR1	REST corepressor 1	1	0	1	0	1	0	1	0
FGF7	fibroblast growth factor 7 (keratinocyte growth factor)	1	1	0	0	0	0	0	0
G3BP2	GTPase activating protein (SH3 domain) binding protein 2	1	0	1	0	1	0	1	0
HIVEP2	human immunodeficiency virus type I enhancer binding protein 2	1	1	0	0	0	0	0	0
WEE1	WEE1 homolog (S. pombe)	1	1	0	0	0	0	0	0
DCUN1D3	DCN1, defective in cullin neddylation 1, domain containing 3 (S. cerevisiae)	1	0	0	1	1	0	1	0
WWC1	WW and C2 domain containing 1	1	1	0	0	0	0	0	0
SCG2	secretogranin II (chromogranin C)	1	0	1	0	0	0	0	0
LCORL	ligand dependent nuclear receptor corepressor-like	1	0	1	0	0	0	0	0
SOX11	SRY (sex determining region Y)-box 11	1	0	1	0	0	0	0	0
MAP3K7IP2	mitogen-activated protein kinase kinase 7 interacting protein 2	1	1	0	0	0	0	0	0
LOC153222	adult retina protein	1	0	1	0	0	0	0	0
RPS6KA3	ribosomal protein S6 kinase, 90kDa, polypeptide 3	1	0	1	0	0	0	0	0
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog	1	0	1	0	0	0	0	0
MIER3	mesoderm induction early response 1, family member 3	1	0	1	0	0	0	0	0
KBTBD2	kelch repeat and BTB (POZ) domain containing 2	1	0	1	0	1	0	0	1
SP3	Sp3 transcription factor	1	0	1	0	0	0	0	0
SATB1	SATB homeobox 1	1	0	1	0	0	0	0	0
BRD1	bromodomain containing 1	1	0	1	0	0	0	0	0
ZMYM2	zinc finger, MYM-type 2	1	0	1	0	0	0	0	0
PICALM	phosphatidylinositol binding clathrin assembly protein	2	0	1	1	0	0	0	0
FBXO11	F-box protein 11	1	0	1	0	0	0	0	0
COPS3	COP9 constitutive photomorphogenic homolog subunit 3 (Arabidopsis)	1	0	0	1	0	0	0	0
CAB39	calcium binding protein 39	1	1	0	0	0	0	0	0
NOVA1	neuro-oncological ventral antigen 1	1	0	1	0	0	0	0	0
MYB	v-myb myeloblastosis viral oncogene homolog (avian)	1	0	1	0	1	0	1	0
PLAG1	pleiomorphic adenoma gene 1	1	0	0	1	1	0	1	0
BDNF	brain-derived neurotrophic factor	1	0	0	1	0	0	0	0
RKHD3	ring finger and KH domain containing 3	1	0	0	1	0	0	0	0
KCNA1	potassium voltage-gated channel, shaker-related subfamily, member 1 (episodic ataxia with myokymia)	1	0	1	0	0	0	0	0
MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	1	0	0	1	0	0	0	0
YWHAE	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide	1	0	1	0	0	0	0	0
ZNF644	zinc finger protein 644	1	0	0	1	0	0	0	0
CAMTA1	calmodulin binding transcription activator 1	1	0	1	0	0	0	0	0
CHD7	chromodomain helicase DNA binding protein 7	1	0	1	0	0	0	0	0
VT11A	vesicle transport through interaction with t-SNAREs homolog 1A (yeast)	1	0	0	1	0	0	0	0
SOX10	SRY (sex determining region Y)-box 10	1	0	0	1	0	0	0	0
AZIN1	antizyme inhibitor 1	1	0	0	1	1	0	0	1
NR2F2	nuclear receptor subfamily 2, group F, member 2	1	0	0	1	0	0	0	0
KIAA1267	KIAA1267	1	0	0	1	0	0	0	0
PAPOLA	poly(A) polymerase alpha	1	0	0	1	0	0	0	0
RICTOR	rapamycin-insensitive companion of mTOR	1	0	0	1	0	0	0	0
CACNA1C	calcium channel, voltage-dependent, L type, alpha 1C subunit	1	0	0	1	0	0	0	0
C10orf12	chromosome 10 open reading frame 12	1	0	0	1	0	0	0	0
FUBP1	far upstream element (FUSE) binding protein 1	1	0	0	1	0	0	0	0
CKAP5	cytoskeleton associated protein 5	1	0	0	1	0	0	0	0
RNF111	ring finger protein 111	1	0	0	1	0	0	0	0
ELAVL2	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2 (Hu antigen B)	1	0	0	1	0	0	0	0
ZNF703	zinc finger protein 703	1	0	0	1	0	0	0	0
ASTN2	astrotactin 2	1	0	0	1	0	0	0	0
KPNA4	karyopherin alpha 4 (importin alpha 3)	1	0	0	1	0	0	0	0
CSNK1A1	casein kinase 1, alpha 1	1	0	0	1	0	0	0	0
FRYL	FRY-like	1	0	0	1	0	0	0	0
SERTAD2	SERTA domain containing 2	1	0	0	1	0	0	0	0
CTDSP12	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase like 2	1	0	0	1	0	0	0	0
FLJ20273	RNA-binding protein	1	0	0	1	0	0	0	0
SOCS1	suppressor of cytokine signaling 1	1	1	0	0	0	0	0	0
MLSTD2	male sterility domain containing 2	1	0	1	0	0	0	0	0

Figure 46 : Prédiction des cibles de gga-miR-155 et de son orthologue viral mdv1-miR-M4-5P au sein des 3'UTR de *gallus gallus* par le logiciel TargetScan custom.

Les cibles soulignées en bleu ont été validées pour le miARN humain hsa-miR-155, les cibles soulignées en vert ont été validées pour le miARN gga-miR-155 et son orthologue mdv1-miR-M4-5P et les cibles choisies pour notre étude sont soulignées en rouge.

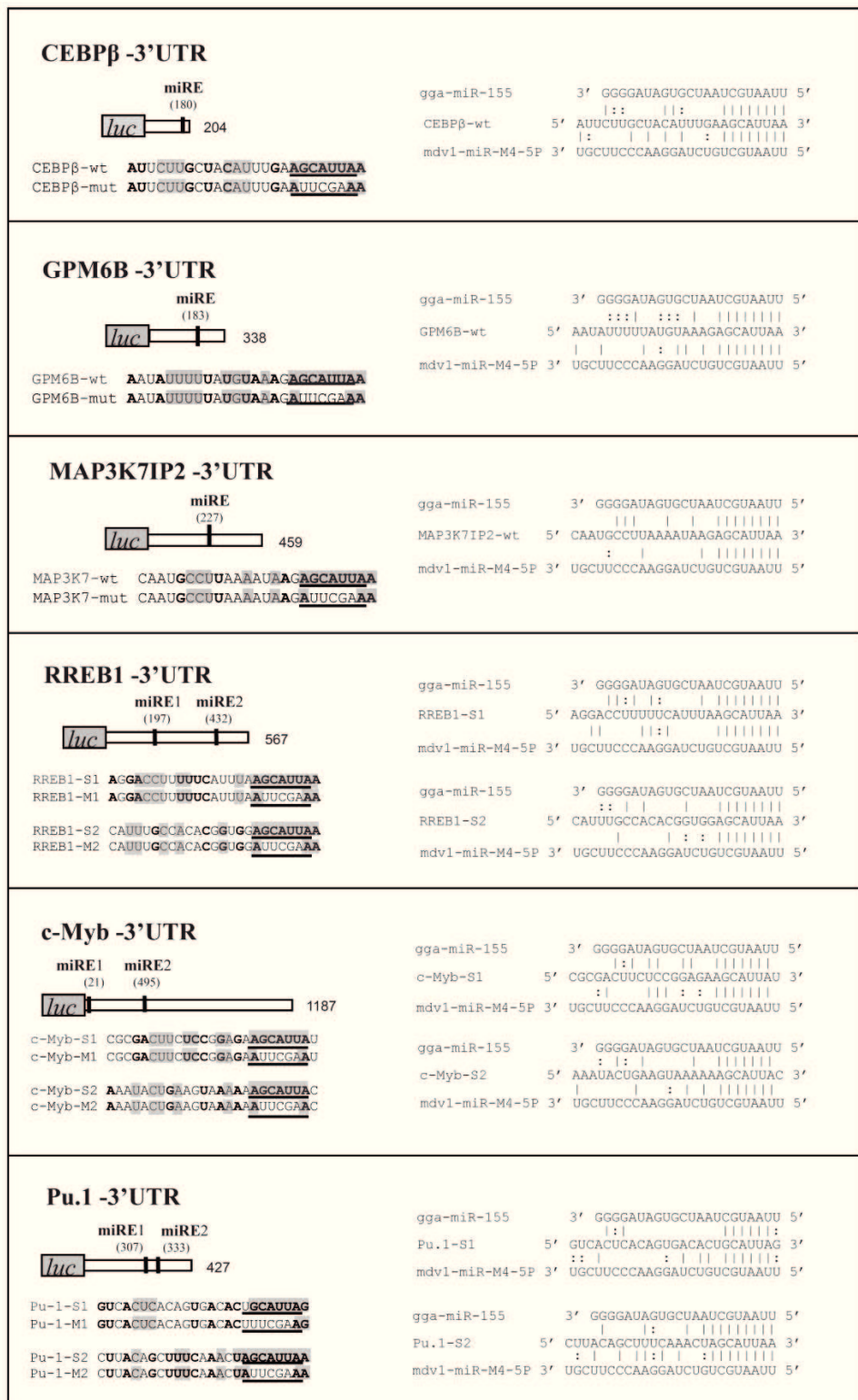


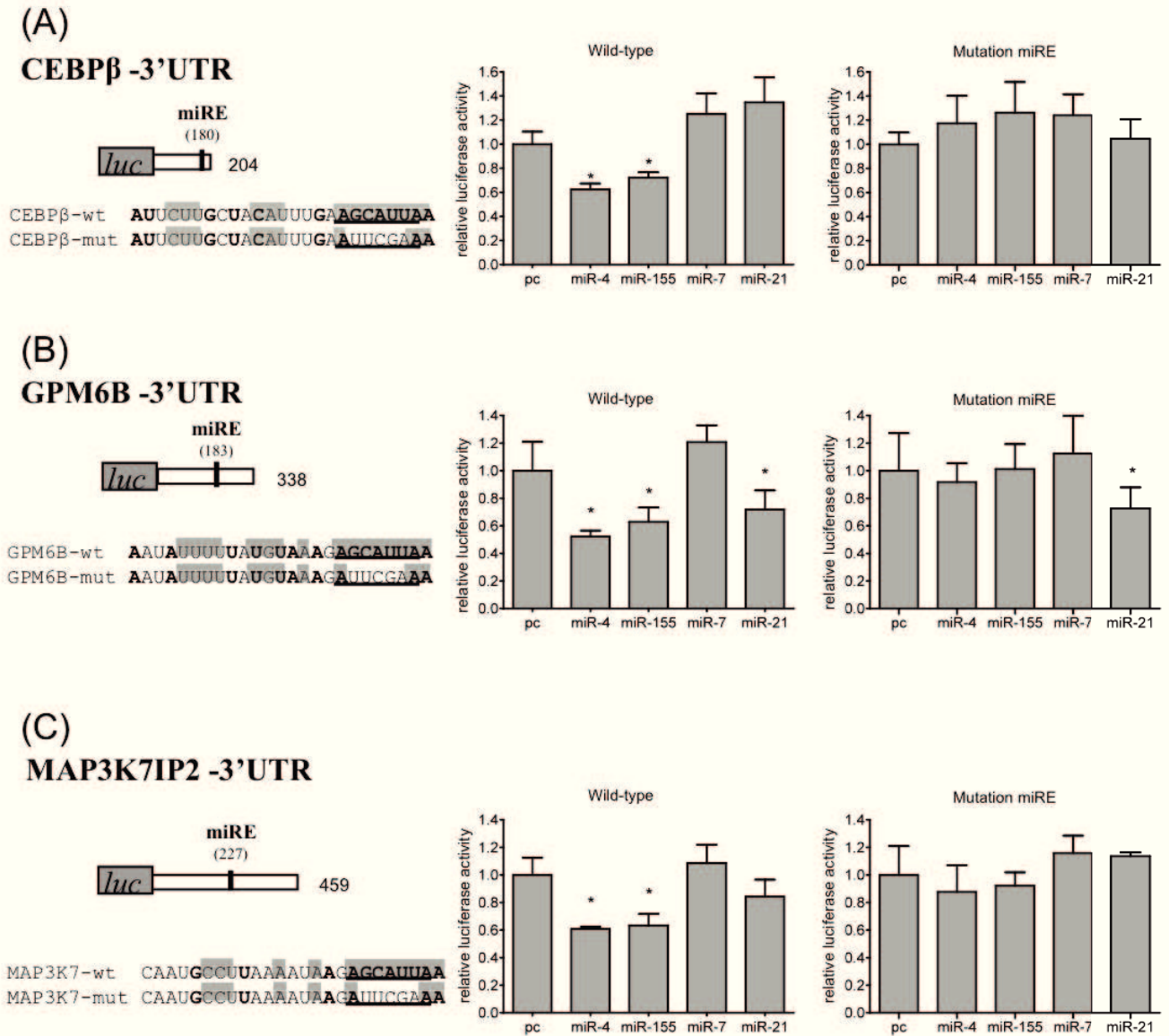
Figure 47 : Localisation schématique des miRE au sein de la région 3'UTR de six gènes cellulaires cibles. Pour chaque construction, la taille de la fraction 3'UTR clonée en aval de la luciférase *Renilla* (*luc*), ainsi que la position de chaque miRE sont indiquées en haut à gauche. En dessous, sont reportées les séquences de chaque miRE sauvage ainsi que de leur forme mutée, la séquence seed correspondante est soulignée, les bases complémentaires à mdv1-miR-M4 sont représentées en gras et les bases complémentaires à gga-miR-155 sont surlignées en gris. L'alignement entre chaque miRE et les miARN mdv1-miR-M4-5P et gga-miR-155 est présenté dans la partie droite de chaque cartouche.

viral oncogène homologue) et *RREB1* (ras responsive element binding protein 1)). Au sein de ces cibles, *CEBPβ* et *PU.1* avaient été validées auparavant pour mdv1-miR-M4 et gga-miR-155 (Zhao *et al.*, 2009b), tandis que le miRE de *MAP3K7IP2* avait uniquement été validé dans le modèle humain (Ceppi *et al.*, 2009).

1.3.1. Evaluation de la fonctionnalité des miRE lors d'expression ectopique de mdv1-pré-miR-M4 et gga-pré-miR-155.

Un large segment des 3'UTR de chacun des gènes cibles potentiels renfermant les miRE a été inséré en aval du gène rapporteur luciférase *Renilla* au sein du vecteur pRL-TK pour donner les constructions pRL-TK-CEBPβ-wt, pRL-TK-GPM6B-wt, pRL-TK-MAP3K7IP2-wt, pRL-TK-RREB1-wt, pRL-TK-c-Myb-wt et pRL-TK-PU.1-wt. Puis, la séquence des miRE complémentaire des séquences « seed » des miARN a été modifiée spécifiquement par mutagenèse dirigée (Figure 47), aboutissant aux constructions pRL-TK-CEBPβ-mut, pRL-TK-GPM6B-mut et pRL-TK-MAP3K7IP2-mut pour les 3'UTR ne renfermant qu'un seul miRE potentiel. Enfin, pour les 3' UTR renfermant deux miRE potentiels, trois constructions mutantes ont été effectuées systématiquement correspondant aux miRE mutées séparément ou à la double mutation (pRL-TK-RREB1-M1, pRL-TK-RREB1-M2 et pRL-TK-RREB1-M1M2 ; pRL-TK-c-Myb-M1, pRL-TK-c-Myb-M2 et pRL-TK-c-Myb-M1M2 ; pRL-TK-PU.1-M1, pRL-TK-PU.1-M2 et pRL-TK-PU.1-M1M2).

Afin d'évaluer l'effet des deux miARN sur les différentes 3'UTR, l'ensemble des constructions PRL-TK natives et mutées ont été co-transfectées dans des cellules DF-1 avec les vecteurs pcDNA3.1 sans insert, pcDNA-miR-M4, pcDNA-miR-155, pcDNA-miR-M7 et pcDNA-miR-21 exprimant respectivement mdv1-pré-miR-M4, gga-pré-miR-155 ou les témoins mdv1-pré-miR-M7 et gga-pré-miR-21, ainsi qu'un vecteur pcDNA-Mluc exprimant la luciférase firefly dans le but de normaliser chaque transfection. Afin de pouvoir comparer les différents essais, les valeurs d'activité luciférase correspondant à chaque construction de cibles ont été standardisées en les rapportant aux activités mesurées pour le vecteur pcDNA3.1 sans insert. Ainsi, il est apparu que les miRE des trois gènes *CEBPβ*, *GPM6B* et *MAP3K7IP2* possédant un miRE unique pour mdv1-miR-M4 et gga-miR-155 étaient fonctionnels aussi bien avec mdv1-miR-M4-5P qu'avec gga-miR-155 (Figure 48). En effet, l'expression des miARN mdv1-miR-M4 et gga-miR-155 entraînait une diminution significative d'environ 40% d'activité luciférase pour les constructions (PRL-TK-CEBPβ-wt, PRL-TK-GPM6B-wt, PRL-TK-MAP3K7IP2-wt) en comparaison des activités correspondant aux miRE mutés (PRL-TK-CEBPβ-mut, PRL-TK-GPM6B-mut, PRL-TK-MAP3K7IP2-mut; Figure 48). A l'inverse, aucune différence significative n'a été observée dans les mêmes conditions lors de la co-transfection avec les vecteurs pcDNA-miR-M7 ou pcDNA-miR-21 (Figure 48). Par ailleurs, une diminution significative a été observée pour les deux constructions sauvages et mutées contenant la



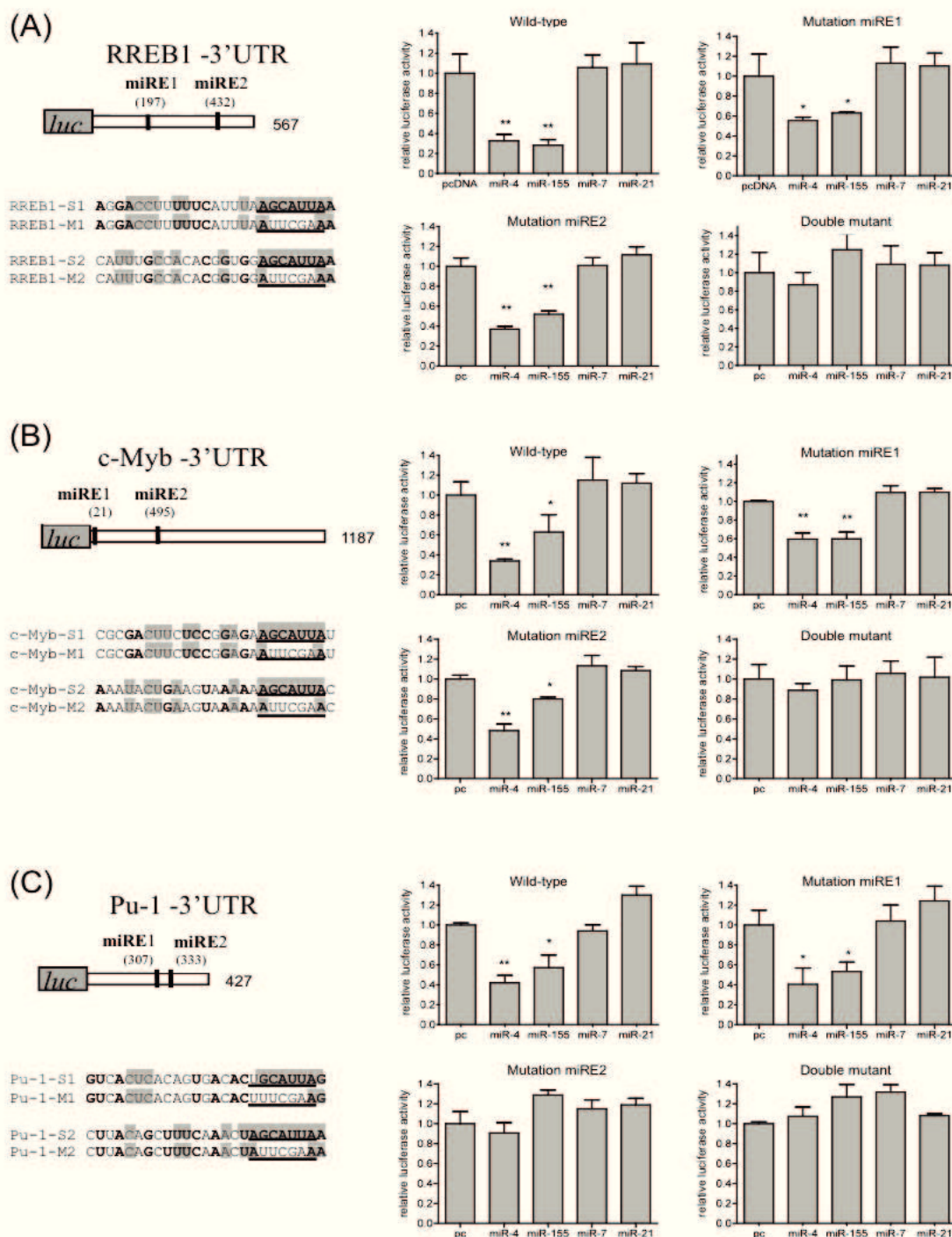


Figure 49: Etude comparative de l'effet des microARN mdv1-miR-M4 (miR-4) et gga-miR-155 (miR-155) sur les cibles possédant deux miRE.

Le détail des constructions présenté dans la figure 47 est reporté dans la partie gauche de la figure. L'effet des microARN mdv1-miR-M4 (miR-4), gga-miR-155 (miR-155), ainsi que des 2 miARN témoins mdv1-miR-M7 (miR-7), et gga-miR-21 (miR-21), sur les cibles cellulaires possédant deux miRE et leurs mutants respectifs (A) RREB1, (B) c-Myb et (C) PU.1 est présenté sous la forme d'histogrammes. Les valeurs luciférase obtenues lors des co-transfections de cellules DF-1 par les constructions sauvages (pRL-TK-RREB1-wt, pRL-TK-c-Myb-wt, pRL-TK-PU.1-wt) et mutées (pRL-TK-RREB1-M1, pRL-TK-RREB1-M2, pRL-TK-RREB1-M1M2, pRL-TK-c-Myb-M1, pRL-TK-c-Myb-M2, pRL-TK-c-Myb-M1M2, pRL-TK-PU.1-M1, pRL-TK-PU.1-M2, pRL-TK-PU.1-M1M2) avec les vecteurs d'expression pcDNA-mdv1-miR-M4, pcDNA-gga-miR-155, pcDNA-mdv1-miR-M7 et pcDNA-gga-miR-21 ont été relativisées en fonction de la valeur correspondant au pcDNA3.1 sans insert (pc). Les barres d'erreurs indiquent la déviation standard correspondant à 3 essais réalisés en triplicate. La significativité de chaque essai en comparaison de la co-transfection avec le pcDNA3.1 sans insert (pc) a été évaluée par un test de *student*, où * et ** correspondent à une valeur de p respectivement <0,05 et <0,001.

3'UTR de *GPM6B* lors de l'expression ectopique de gga-pré-miR-21 (Figure 48B). Cette diminution équivalente que le miRE correspondant à mdv1-miR-M4 ou gga-miR-155 ait été mutée ou non semble indiquer la présence dans cette 3' UTR d'un miRE pour gga-miR-21 non chevauchant directement avec le miRE de gga-miR-155 et mdv1-miR-M4.

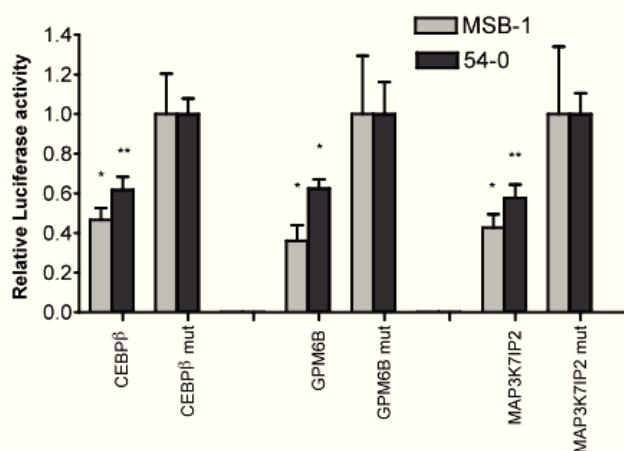
Par ailleurs, les deux pré-miARN (mdv1-pré-miR-M4 et gga-pré-miR-155) entraînaient une répression forte également sur les trois gènes possédant deux miRE (*RREB1*, *c-Myb* et *PU.1*). En effet, la comparaison des activités luciférase correspondant à la construction sauvage avec celles correspondant aux doubles mutants a fait apparaître une forte diminution d'activité luciférase lors de l'expression ectopique des deux miARN orthologues (Figure 49). Alors que l'effet des deux miARN semblait être équivalent sur la construction véhiculant la 3'UTR du gène *RREB1* avec une diminution de l'activité luciférase de l'ordre de 70% (Figure 49A), le miARN mdv1-M4-5P avait un impact plus important que le gga-miR-155 sur celles des deux autres gènes *c-Myb* et *PU.1* (Figure 49B et C). Plus précisément la diminution liée au mdv1-miR-M4 se situait respectivement autour de 70% et de 60% alors que celle correspondant au gga-miR-155 autour de 40% et 45%.

D'autre part, l'analyse des résultats au regard de la fonctionnalité des deux miRE, a montré que seul le miRE2 était fonctionnel au sein de la 3' UTR de *PU.1*, puisque la mutation du miRE1 n'influait pas l'activité luciférase, contrairement à la mutation du miRE2 qui restaurait l'intégralité de l'activité luciférase (Figure 49C). Par contre, pour les gènes *RREB1* et *c-Myb* (Figure 49A et B), les deux miRE semblaient être fonctionnels indépendamment du fait que l'activité luciférase n'ait pas été restaurée lors de la mutation d'une seul des deux miRE. De plus, l'analyse fine des activités luciférase résultantes semblait mettre en évidence un effet coopératif des deux miRE au sein de la 3' UTR de *c-Myb* (Figure 49B) pour le mdv1-miR-M4-5P et au sein de la 3' UTR de *RREB1* pour gga-miR-155 (Figure 49A). Enfin, la comparaison de l'effet entre mdv1-miR-M4-5P et gga-miR-155 sur le site miRE1 de la 3' UTR de *c-Myb* semblait montrer une répression légèrement plus faible pour le miARN cellulaire en comparaison de son orthologue viral (Figure 49B, mutation miRE2).

1.3.2. Evaluation de la fonctionnalité des miRE pour mdv1-miR-M4-5P en système endogène.

Afin d'évaluer la fonctionnalité et l'effet du mdv1-miR-M4-5P endogène sur les différentes cibles cellulaires, les constructions possédant les 3'UTR des différents gènes cellulaires cibles ont été co-transfectées avec le plasmide rapporteur de transfection pcDNA-Mluc, dans des cellules MSB-1 et 54-O qui expriment constitutivement mdv1-miR-M4-5P et dans lesquelles gga-miR-155 n'est pas détectable (Figure 45). Un effet similaire à celui de l'expression ectopique a été observé au sein de ces deux lignées lymphoblastoïdes issues de tumeurs induites par GaHV-2, et ceci pour l'ensemble des 3'UTR contenant un miRE (CEBP β , *GPM6B* et *MAP3K7IP2*, Figure 50A) ou deux miRE (*RREB1*, *c-*

(A)



(B)

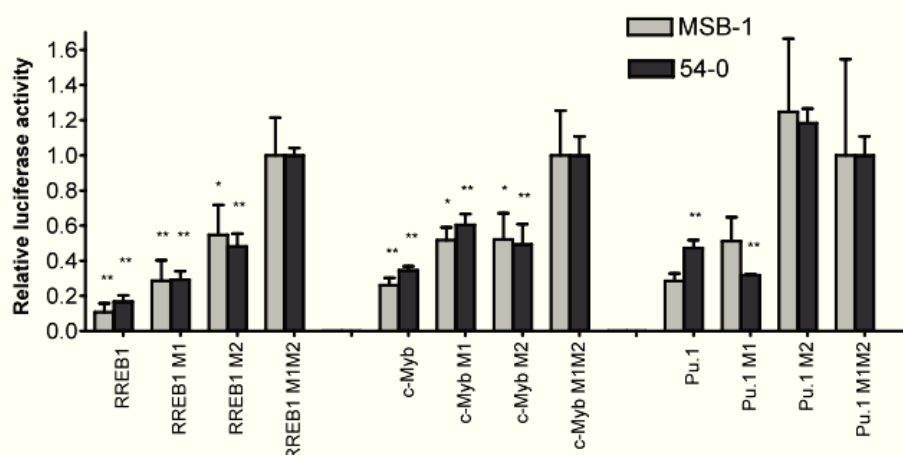


Figure 50 : Effet du microARN mdv1-miR-M4-5P endogène sur les 3'UTR des cibles cellulaires présentant un ou deux miRE.

L'effet du microARN mdv1-miR-M4 sur les cibles cellulaires possédant un miRE et leurs mutants respectifs CEBPβ, GPM6B et MAP3K7IP2 (A) et sur les cibles possédant deux miRE et leurs mutants respectifs RREB1, c-Myb et PU.1 (B) est présenté sous la forme d'histogrammes. Les valeurs luciférase obtenues lors des transfections de cellules 54-O et MSB-1 par les constructions sauvages (pRL-TK-CEBPβ-wt, pRL-TK-GPM6B-wt, pRL-TK-MAP3K7IP2-wt, pRL-TK-RREB1-wt, pRL-TK-c-Myb-wt, pRL-TK-PU.1-wt) et mutées (pRL-TK-RREB1-M1, pRL-TK-RREB1-M2, pRL-TK-c-Myb-M1, pRL-TK-c-Myb-M2, pRL-TK-PU.1-M1, pRL-TK-PU.1-M2) ont été relativisées en fonction de la valeur correspondant à la construction 3'UTR mutante (mut ou M1M2). Les barres d'erreur indiquent la déviation standard correspondant à trois essais réalisés en triplicate. La significativité de chaque essai en comparaison de la co-transfection avec la construction 3'UTR mutante (mut ou M1M2) a été évaluée par un test de *student*, où * et ** correspondent à une valeur de p respectivement <0,05 et <0,001.

Myb et PU.1, Figure 50B). En effet, à l'exception de la construction PRL-TK-RREB1 pour laquelle les taux d'inhibition semblent plus importants lors du test dans les lignées lymphoblastoïdes transformées par GaHV-2, avec un taux d'inhibition de 80% en 54-O et de 90% en MSB1 (Figure 50B) versus 40% en expression ectopique (Figure 49A), les taux d'inhibition observés avec mdv1-miR-M4-5P endogène sont similaires à ceux mesurés lors des tests par expression ectopique de mdv1-pré-miR-M4. De même, il faut noter qu'à l'instar des essais lors des expressions ectopiques, dans ces essais, seul le site miRE2 de la 3'UTR de PU.1 apparaît fonctionnel. Par ailleurs, la comparaison plus fine de l'impact des miRE présents dans la région 3'UTR de RREB1 dans les cellules 54-O et MSB-1 montre que le site miRE2 induirait une inhibition plus importante que le miRE1 (Figure 50B), alors que les deux miRE semblaient présenter un impact équivalent lors d'expression ectopique du mdv1-miR-M4.

1.4. Analyse de l'effet des deux brins du mdv1-pré-miR-M4 et du gga-pré-mir-155 sur des gènes viraux impliqués dans le clivage et l'encapsidation de l'ADN viral.

1.4.1. Recherche bioinformatique de cibles virales potentielles

Compte tenu que les deux miARN mdv1-miR-M4-5P et gga-miR-155 semblaient agir de manière équivalente sur l'expression des six gènes cellulaires, nous avons testé leurs effets sur de potentielles cibles virales. En premier lieu, à l'aide du logiciel BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), nous avons effectué une recherche de séquences partiellement complémentaires des miARN mdv1-miR-M4-5P et mdv1-miR-M4-3P au sein du génome de GaHV-2. Nous avons ainsi défini une cible potentielle au sein de la région codante de UL-28 pour mdv1-miR-M4-5P et au sein de la région codante de UL-32 pour mdv1-miR-M4-3P (Figure 51). Ces deux sites présentent des interactions supplémentaires à la séquence « seed » avec les nucléotides 2-12 de mdv1-miR-M4-5P ciblant une séquence continue de 11 nucléotides de la partie 3' de la région codante de UL-28, et les nucléotides 1-12 de mdv1-miR-M4-3P ciblant une séquence continue de 12 nucléotides de la partie 3' de la région codante de UL-32 (Figure 51).

1.4.2. Impact du mdv1-pré-mir-M4 endogène sur les 2 cibles virales

Afin de tester l'effet des deux brins de mdv1-pré-miR-M4 sur les cibles virales UL-28 et UL-32, les 500 nucléotides entourant les miRE ont été insérés en aval du gène rapporteur luciférase *Renilla* au sein du vecteur pRL-TK sous la forme native ou mutée (Figure 51). Ainsi, quatre

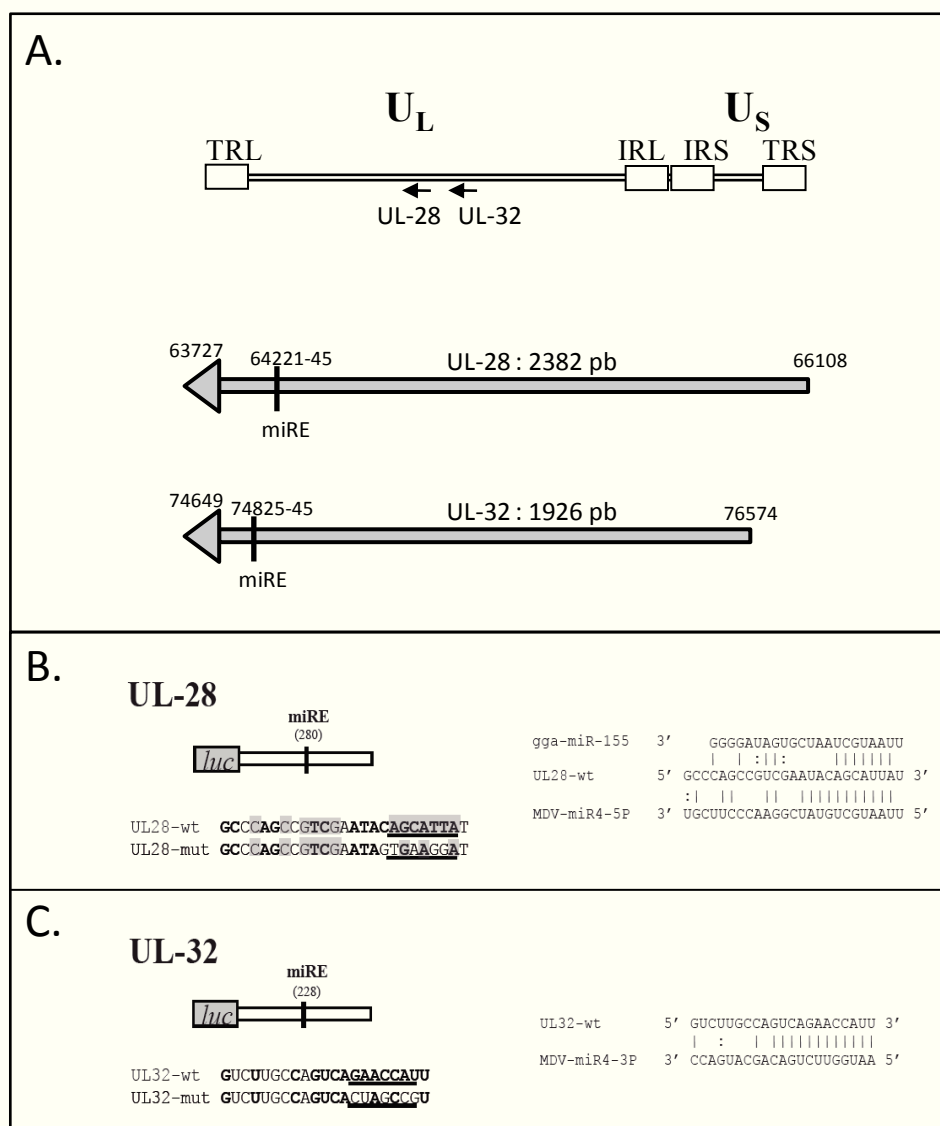


Figure 51: Localisation schématique des miRE de mdv1-miR-M4 présents au sein des régions codantes des gènes viraux UL-28 et UL-32.

(A) Localisation des gènes UL-28 et UL-32 et les miRE de mdv1-pré-miR-M4 présents au sein de leurs séquences codantes sur le génome de GaHV-2 Md5. Les 500 pb entourant les miRE sauvages ou mutés des cibles viraux UL-28 (B) et UL-32 (C) ont été insérées en aval de la luciférase *Renilla* (*luc*) pour former les vecteurs pRL-TK-UL-28-wt, pRL-TK-UL-28-mut, pRL-TK-UL-32-wt et pRL-TK-UL-32-mut. La séquence de chaque miRE sauvage ainsi que de sa forme mutée est indiquée en dessous ; la séquence complémentaire à la séquence « seed » est soulignée, les bases complémentaires à mdv1-miR-M4 sont représentées en gras et les bases complémentaires à gga-miR-155 sont surlignées en gris. L'alignement entre le miRE de UL-28 et les miARN mdv1-miR-M4-5P, gga-miR-155 ainsi que l'alignement entre le miRE d'UL-32 et le miARN mdv1-miR-M4-3P sont présentés dans la partie droite des 2 cartouches.

constructions ont été réalisées, les deux vecteurs sauvages pRL-TK-UL-28-wt et pRL-TK-UL-32-wt et les deux vecteurs mutants correspondants (pRL-TK-UL-28-mut et pRL-TK-UL-32-mut).

Dans un premier temps, l'effet des deux miARN mdv1-miR-M4-3P et 5P sur ces cibles potentielles a été évalué lors de co-transfection de pcDNA-Mluc avec les différentes constructions sauvages et mutées dans les lignées lymphoblastoïdes MSB-1 et 54-O exprimant de manière constitutive les deux brins du mdv1-pré-miR-M4. Ainsi il est apparu que les deux miARN endogènes exerçaient un effet sur les cibles des deux gènes viraux testés (Figure 52A). Plus précisément, il est apparu que le miRE de la cible UL28 était particulièrement bien reconnu par mdv1-miR-M4-5P, puisque ce dernier induisait une réduction de l'activité luciférase de l'ordre de 80% et 90% respectivement dans les deux lignées MSB-1 et 54-O. Cependant, pour la cible virale UL-32, bien que la complémentarité entre son miRE et le mdv1-miR-M4-3P soit équivalente à celle de la cible UL-28 avec le mdv1-miR-M4-5P, la diminution d'activité luciférase n'est que de l'ordre de 50% (Figure 52A). Cette diminution moins importante peut s'expliquer par la plus faible proportion de l'expression de mdv1-miR-M4-3P par rapport à celle de mdv1-miR-M4-5P observée au sein des cellules MSB-1 et 54-O en northern blot (Figure 45) et lors des différentes banques réalisées au laboratoire.

1.4.3. Impact du mdv1-pré-miR-M4 et du gga-pré-miR-155 sur les 2 cibles virales lors d'expressions ectopiques.

L'interaction entre les deux brins des mdv1-pré-miR-M4 et gga-pré-miR-155 et les cibles virales UL-28 et UL-32 a ensuite été testée lors de co-transfections tripartites dans la lignée cellulaire fibroblastique DF-1 et dans la lignée lymphoblastoïde DT40 des vecteurs pcDNA3.1 sans insert, pcDNA-miR-M4, pcDNA-miR-155 et pcDNA-miR-21 exprimant respectivement mdv1-pré-miR-M4, gga-pré-miR-155 ou le témoin gga-pré-miR-21, avec chaque construction pRL-TK-UL-28-wt, pRL-TK-UL-32-wt, pRL-TK-UL-28-mut et pRL-TK-UL-32-mut, ainsi que le vecteur pcDNAM-luc exprimant la luciférase firefly permettant de normaliser chaque transfection. Les résultats obtenus montraient une baisse d'activité luciférase de 60% dans la lignée DF-1 et d'environ 40 % dans la lignée DT-40, pouvant être reliée à la présence du miRE de mdv1-miR-M4-5P au sein d'UL28 (Figure 52B). Par contre dans ces conditions d'expression ectopique, avec un taux de transfection de l'ordre de 50 et 20% respectivement dans les cellules DF-1 et DT-40, ainsi qu'une expression potentiellement faible du brin 3P à partir du pré-miARN, l'effet de mdv1-miR-M4-3P sur UL32 n'a pu être mis en évidence puisqu'il n'y a pas eu de réduction d'activité luciférase en cellules DT40 et que celle observée en cellules DF-1 de 30% était non significative.

Enfin, à l'image du témoin négatif gga-miR-21, l'expression ectopique du gga-pré-miR-155 ne semblait pas avoir d'effet sur l'expression des deux gènes viraux UL-28 et UL-32 (Figure 52B).

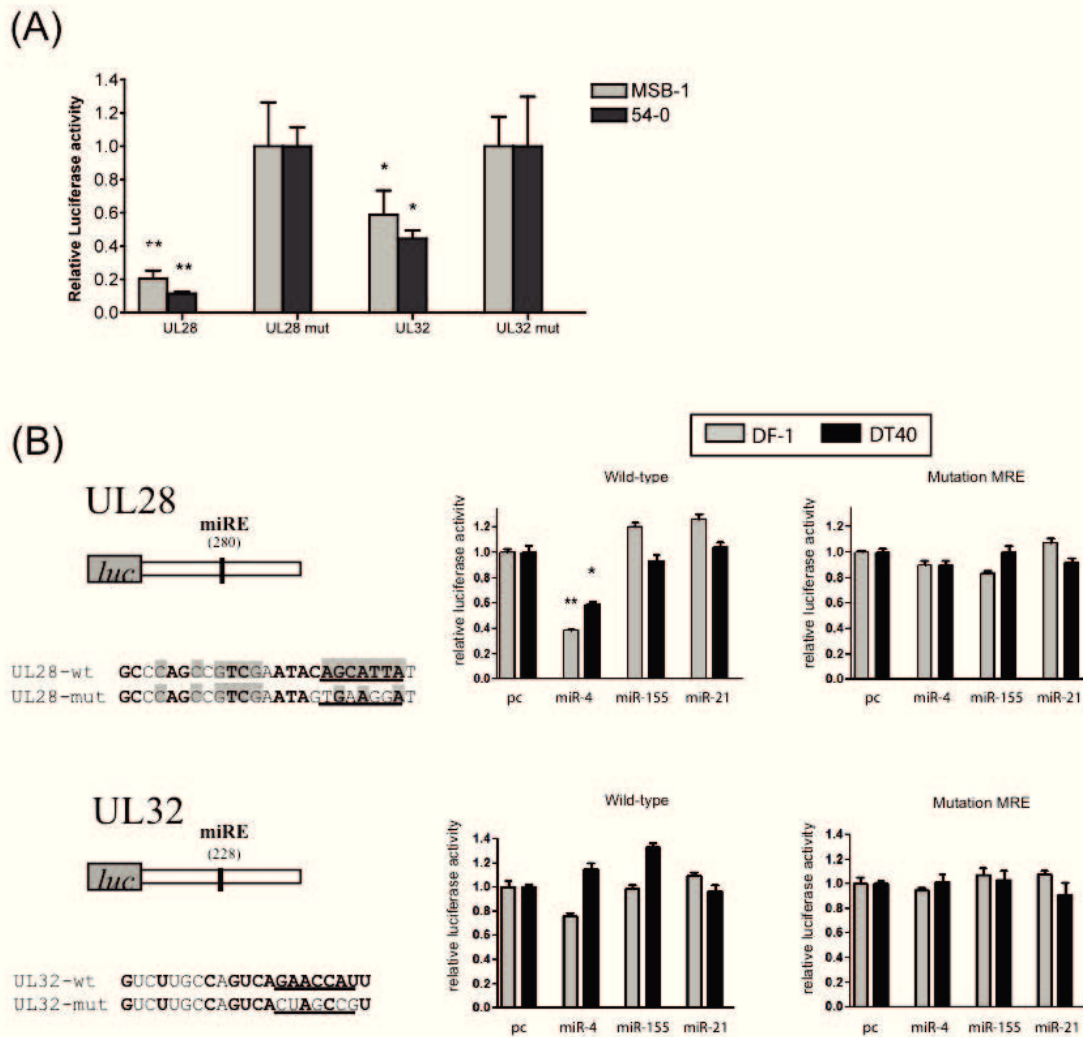


Figure 52 : Effet du pré-miARN mdv1-miR-M4 sur les miRE des cibles virales UL-28 et UL-32.

(A) Effet du pré-miARN mdv1-pré-miR-M4 endogène sur les miRE des cibles virales UL-28 et UL-32. L'effet du microARN mdv1-miR-M4 sur les cibles virales UL-28 et UL-32 et leurs mutants respectifs UL-28mut et UL-32mut est présenté sous la forme d'histogrammes. Les valeurs luciférase obtenues lors des transfections de cellules 54-O et MSB1 par les constructions sauvages (pRL-TK-UL28-wt et pRL-UL32-wt) ont été relativisées en fonction de la valeur correspondant à la construction 3'UTR mutante (pRL-TK-UL28-mut, pRL-TK-UL32-mut).

(B) Effet de la surexpression en cellules DF-1, des microARN mdv1-miR-M4 (miR-4), gga-miR-155 (miR-155) sur les cibles virales UL-28 et UL-32 et leurs mutants respectifs. Le détail des constructions présenté dans la figure 51 est reporté dans la partie gauche de la figure. L'effet des microARN mdv1-miR-M4 (miR-4), gga-miR-155 (miR-155), ainsi que du miARN témoin gga-miR-21 (miR-21) sur les cibles virales UL-28 et UL-32 et leurs mutants respectifs UL-28mut et UL-32mut est présenté sous la forme d'histogrammes. Les valeurs luciférase obtenues lors des co-transfections de cellules DF-1 par les constructions sauvages et mutées avec les vecteurs d'expression pcDNA-mdv1-miR-M4, pcDNA-gga-miR-155 et pcDNA-mdv1-miR-M7 ont été relativisées en fonction de la valeur correspondant au pcDNA3.1 sans insert (pc). Les barres d'erreurs indiquent la déviation standard correspondant à trois essais réalisés en triplicate. La significativité de chaque essai en comparaison de la transfection avec la construction 3'UTR mutante (A) ou du témoin pcDNA3.1 sans insert (B) a été évaluée par un test de *student*, où * et ** correspondent à une valeur de p respectivement <0,05 et <0,001.

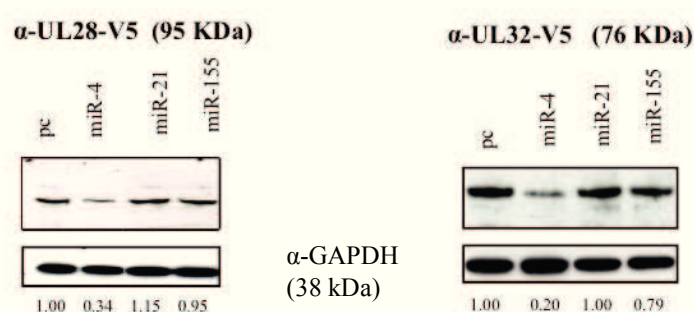


Figure 53 : Evaluation par Western blot de l'effet du mdv1-pré-miR-M4 sur la production des protéines UL-28 et UL-32.

Les lysats de cellules LMH co-transfectées par les constructions pcDNA-UL-28/V5 ou pcDNA-UL-32/V5 et les vecteurs d'expression pcDNA3.1 sans insert, pcDNA-mdv1-miR-M4, pcDNA-gga-miR-21 ou pcDNA-gga-miR-155, ont été déposées sur un gel SDS-PAGE (7,5%). L'immunoblot a été réalisé avec un anticorps monoclonal de souris dirigé contre l'épitope V5 et un anticorps monoclonal de souris anti-GAPDH, puis avec des anticorps secondaires de lapin anti souris couplés à la peroxydase. L'intensité de signal observée pour les protéines UL-28-V5 et UL-32-V5 ont été rapportées à l'intensité du signal obtenu pour GAPDH à l'aide du logiciel Image J (Wayne Rasband, NIH image, USA). Les valeurs obtenues ont été relativisées par celles de la co-transfection avec le vecteur pcDNA3.1 sans insert (pc) et sont indiquées en dessous de chaque dépôt. Des résultats similaires ont été obtenus lors de trois essais de co-transfection indépendants et les données de l'essai présenté sont représentatives.

1.4.4. Effet des mdv1-pré-miR-M4 et du gga-pré-miR-155 sur l'expression des protéines UL 28 et UL-32

Comme tous les essais précédents avaient été réalisés sur la base de systèmes rapporteurs luciférase dans lesquels la région miRE naturellement située à l'intérieur du cadre de lecture avaient été clonée au sein de la 3'UTR du gène de la luciférase, nous avons voulu visualiser par immunoblot l'impact des pré-miARN mdv1-pré-miR-M4 et gga-pré-miR-155 sur l'expression des protéines UL-28 et UL-32. Comme nous ne disposions pas d'anticorps contre les protéines UL-28 et UL-32 de GaHV-2, les gènes de ces deux protéines ont été clonés dans le vecteur pcDNA3.1/V5&His. Ainsi, les constructions résultantes pcDNA3.1-UL-28-V5 et pcDNA3.1-UL-32-V5 permettaient l'expression des protéines UL-28-V5 de 95 KDa et UL-32-V5 de 76 KDa, correspondant respectivement aux protéines UL-28 et UL-32 taguées avec l'épitope V5 à leur extrémité C-terminale. L'expression des protéines a été évaluée par immunoblot en comparaison de l'expression de GAPDH, réalisé à partir d'extraits protéiques de cellules LMH préalablement co-transfectées d'une part avec les vecteurs pcDNA3.1-UL-28-V5 ou pcDNA3.1-UL-32-V5 et avec les vecteurs d'expression des pré-miARN (pcDNA-pré-miR-M4, pcDNA-pré-miR-155 ou pcDNA-pré-miR-21) ou le vecteur pcDNA3.1 sans insert, d'autre part. Ces essais avec les miRE dans leur contexte natif ont parfaitement confirmé les résultats précédents. En effet, dans ces conditions expérimentales, il s'est avéré que le mdv1-pré-miR-M4 diminuait fortement l'expression des deux protéines taguées avec une inhibition de 65% de l'expression de la protéine UL-28-V5 et de 80% de celle de la protéine UL-32-V5 (Figure 53). De même, à l'instar de l'effet de l'expression ectopique des miARN en système rapporteur luciférase dans les cellules DF-1 et DT-40, aucune différence d'expression de la protéine UL-28-V5 n'a été observée lors de la co-transfection avec les vecteurs pcDNA-miR-155, pcDNA-miR-21 et pcDNA3.1 sans insert (Figure 53).

2. Etude transcriptionnelle du brin codant les miARN présent sur l'IR_L de GaHV-2.

2.1. Recherche des transcrits codant les miARN présents au sein de l'IR_L de GaHV-2 en latence dans la lignée lymphoïde MSB1.

Quelle que soit la phase du cycle viral, réplication lytique ou latence, les miARN de GaHV-2 présentent des taux d'expression différents les uns des autres, même s'ils sont localisés au sein d'un même cluster. Notamment, notre équipe a récemment montré qu'au sein du cluster1, les miARN mdv1-miR-M4-5P et M3-5P sont retrouvés les plus exprimés dans l'ensemble des quatre banques de miARN réalisées par séquençage Solexa, à partir d'ARN extraits de pulpes de plumes ou de PBL prélevés à 4 semaines pi sur des poulets infectés par GaHV-2 RB-1B ou GaHV-2 CVI988 (Tableau 16). Plus précisément, mdv1-miR-M4-5P représentait 70 à 85% des miARN du cluster 1 et mdv1-miR-M3-5P entre 7 et 16%. Les miARN mdv1-miR-M12-3P, M2-3P et M1-5P représentaient chacun entre 1 et 5% des miARN de ce cluster tandis que le taux relatif des autres miARN du cluster1 qui ont tous été détectés par cette approche était inférieur à 0,1%. Il est intéressant de noter qu'à ce stade tardif de l'infection, les taux d'expression des miARN au sein de ce cluster dans les cellules de pulpe de plumes, siège d'une réplication productive de GaHV-2, étaient équivalents à ceux trouvés dans les PBL correspondant essentiellement à des cellules tumorales dans lesquelles le virus est en latence. Par ailleurs, une différence de taux d'expression de certains miARN du cluster1 lors d'une infection de CEF *in vitro* par la souche GaHV-2 RB-1B avait également été décrite précédemment (Burnside *et al.*, 2006 ; Burnside *et al.*, 2008). Cependant le profil d'expression des miARN était différent de celui que notre laboratoire a obtenu *in vivo*. En effet, même si lors de l'infection répliquative dans ce système, le mdv1-miR-M4-5P était toujours fortement exprimé, les miARN mdv1-miR-M2-5P et M1-5P apparaissaient également fortement exprimés, avec un taux relatif compris entre 22 et 37% des miARN du cluster1 (Tableau 16). Par contre, le taux d'expression des autres miARN de ce cluster était identique à celui que notre équipe a observé lors de l'analyse des quatre banques, notamment le miARN mdv1-miR-M5-3P représentant 7 à 10% des miARN du cluster et les autres miARN représentant moins de 3% des miARN viraux.

Aussi, sachant que des profils d'épissage complexes avaient été décrits dans la partie 5' de la région IR_L/TR_L, j'ai caractérisé en premier lieu, la carte des transcrits alternatifs de toute la région et surtout de la région recouvrant le cluster1 de miARN dans le but de rechercher si l'épissage alternatif et/ou l'utilisation de promoteurs alternatifs pouvaient être impliqués dans l'expression différentielle des miARN de ce cluster. Pour cette étude, j'ai utilisé deux approches complémentaires, la RACE-PCR et la génération de banques de fragments de RT-PCR. Afin d'établir ces dernières, les cDNA correspondant aux ARN totaux extraits de la lignée lymphoblastoïde MSB-1 ont été synthétisés à partir d'un mélange d'amorces (oligonucléotide dT et hexamères aléatoires). Après 25 cycles de PCR, les produits d'amplification obtenus à partir de différents couples d'amorce ont été clonés dans le vecteur pGEMT-easy directement à partir du mélange de PCR, afin d'assurer un maximum d'exhaustivité et d'éviter la sélection d'amplicons particuliers. Enfin, le séquençage d'au moins 20 clones a été effectué pour chaque PCR.

	Pulpes de plumes infectées RB-1B §	Pulpes de plumes infectées CVI988 §	PBL infectées RB-1B §	PBL infectées CVI988 §	CEF infectées RB-1B ‡	CEF infectées RB-1B ‡
mdv1-miR-M5-5P #	0 *	0	0	0	0	1,5
mdv1-miR-M5-3P	1,6	0,5	0,3	1,0	9,7	7,5
mdv1-miR-M12-3P	5,4 ⁰	1,2	1,5	3,3	0	0,4
mdv1-miR-M3-5P	16,0	14,8	6,7	8,2	3,2	1,6
mdv1-miR-M2-3P	5,2	3,9	1,1	3,3	4	1,9
mdv1-miR-M2-5P	0,9	0,5	1,9	0,7	22,6	23,1
mdv1-miR-M4-5P	68,3	75,5	85,1	78,6	32,3	24,9
mdv1-miR-M4-3P	0,6	0,3	0,7	1,1	2,4	0,7
mdv1-miR-M11-5p	0,1	0,6	0,1	0,0	0	0
mdv1-miR-M11-3p	0	0	0	0	0	1,2
mdv1-miR-M1-5P	1,6	2,0	2,2	3,3	25,8	36,7

Tableau 16 : Taux d'expression des microARN du cluster 1 de GaHV-2.

§ Quatre banques de miARN réalisées au laboratoire par séquençage Solexa,

‡,† deux banques de miARN de CEF infectées par la souche RB-1B réalisées par Burnside et ses collaborateurs (‡ Burnside et al., 2006 ; † Burnside et al., 2008).

miARN du cluster 1 dont le taux de détection représentait plus de 0,1% de la totalité des miARN de ce cluster, dans l'une des banques présentées

* le pourcentage indiqué pour chaque miARN correspond au taux de représentativité de ce miARN au sein du cluster1,

⁰ les miARN dont le taux était supérieur à 5% sont indiqués en rouge.

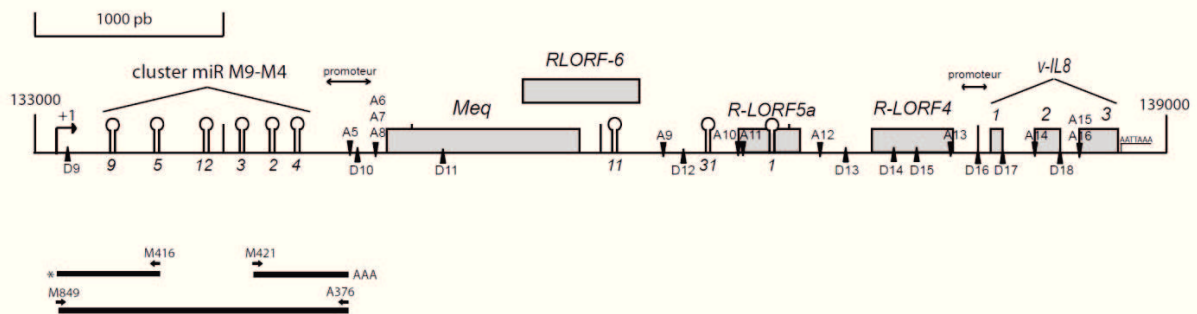


Figure 54: Représentation schématique du transcrit non épissé encodant les miARN mdv1-miR-M9-M4 du cluster 1 de miARN de GaHV-2 en cellules MSB-1.

La région comprise entre les bases 133000 et 139000 de GaHV-2 Md5 est représentée dans la partie supérieure du schéma. Les ORF sont indiquées par des rectangles gris et les miARN par des tiges boucles. Les bornes d'épissage accepteurs sont schématisées par des triangles noirs inversés et les bornes d'épissage donneuses par des triangles noirs.

La localisation des fragments obtenus par RACE-PCR 5' et 3' ainsi que de la PCR spécifique de l'ARNm des miARN sont indiquées en dessous. Les extrémités 5' et 3' du transcrit non épissé ont été identifiées par 5' et 3' RACE PCR respectivement à l'aide des oligonucléotides M416 et M421. Le transcrit non épissé a été amplifié par RT-PCR à l'aide du couple d'amorce M849/A376. Les amorces utilisées pour chaque amplification sont représentées, l'étoile (*) correspond à l'extrémité 5' identifiée par 5'RACE PCR et la queue polyadénylée identifiée par 3'RACE PCR est symbolisée par AAA.

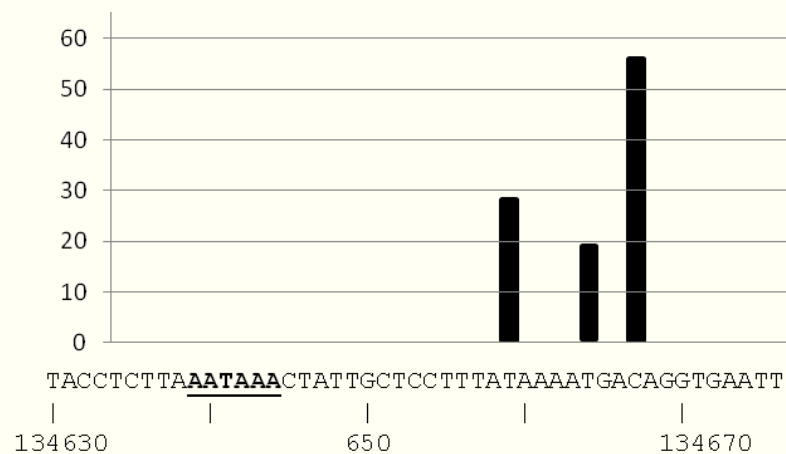


Figure 55: Localisation de l'extrémité 3' polyadénylée du transcrit recouvrant les miARN mdv1-miR-M9-M4.

Une banque de cDNA correspondant à la 3' RACE-PCR initiée par l'amorce M421 à partir des ARN totaux de cellules MSB-1 a été obtenue par clonage des amplicons (sans purification) dans le vecteur pGEM-T-easy. Les barres localisent les 3 extrémités d'ARNm rencontrées au sein des banques sur la séquence comprise entre les bases 134630 et 134676 de la souche GaHV-2 Md5 (donnée en abscisse) et donnent le pourcentage relatif de chacune de ces extrémités. Le signal de polyadénylation est représenté en caractère gras et est souligné.

2.1.1. Recherche d'un ARNm spécifique des mdv1-miR-M9 à mdv1-miR-M4.

La recherche des extrémités 5' et 3' du transcrit codant les miARN a été réalisée par RACE-PCR respectivement à l'aide de l'oligonucléotide reverse M416 situé au niveau du miARN mdv1-miR-M5 et de l'oligonucléotide M421 situé entre mdv1-miR-M3 et mdv1-miR-M2 (Figure 54). Ainsi, un site d'initiation de la transcription (TSS : Transcriptional Start Site) unique en position 133117 du génome de GaHV-2 Md5 a été identifié. Par contre trois extrémités polyadénylées ont été recensées en position 134659, 134664 et 134667 soit respectivement 15, 20 et 23 nucléotides après la fin du signal de polyadénylation AATAAA (Figure 55). Ensuite, une RT-PCR avec les amorces oligonucléotidiques « sens » M849 et « anti-sens » A376, situées respectivement au niveau du TSS précédemment identifié et juste en aval du signal de polyadénylation a permis l'amplification d'un transcrit unique d'une longueur de 1359 pb, dont la séquence a confirmé l'existence d'un ARNm spécifique recouvrant les pré-miR-M9, M5, M12, M3, M2 et M4 (figure 54).

2.1.2. Recherche de transcrits recouvrant l'ensemble des miARN du cluster 1.

Suite à l'identification d'un ARNm renfermant le cluster de pré-miR-M9-M4 et connaissant la complexité de l'épissage alternatif dans cette région du génome viral, nous avons recherché l'existence de transcrits alternatifs renfermant potentiellement les trois autres pré-miR-M1, M11 et M31 situés plus en aval entre le gène *meq* et *R-LORF4*.

Ainsi, une série de RT-PCR a été réalisée en utilisant l'oligonucléotide « sens » M849 localisé au niveau du TSS précédemment identifié et différents oligonucléotides « anti-sens », localisés au niveau du promoteur du gène *meq* (M419), à l'intérieur du cadre de lecture de ce dernier (M420), à l'intérieur du cadre de lecture de *R-LORF5a* (M825), et de *R-LORF4* (M826 ; M827) ou dans l'exon2 de *vIL-8* (M04) (Figure 56A).

Tout d'abord, afin de confirmer l'extrémité 3' du messenger précédent, j'ai réalisé des RT-PCR avec les couples d'amorces M849/M419 et M849/M420. Celles-ci ont permis de mettre en évidence des transcrits alternatifs correspondant à des événements d'épissage situés en aval du signal de polyadénylation précédemment identifié. En effet, de petits transcrits ARNa de 198 pb (avec M419) et ARNb1 à b4 de 270pb, 283pb, 287pb ou 316pb (avec M420) selon l'épissage alternatif ont été

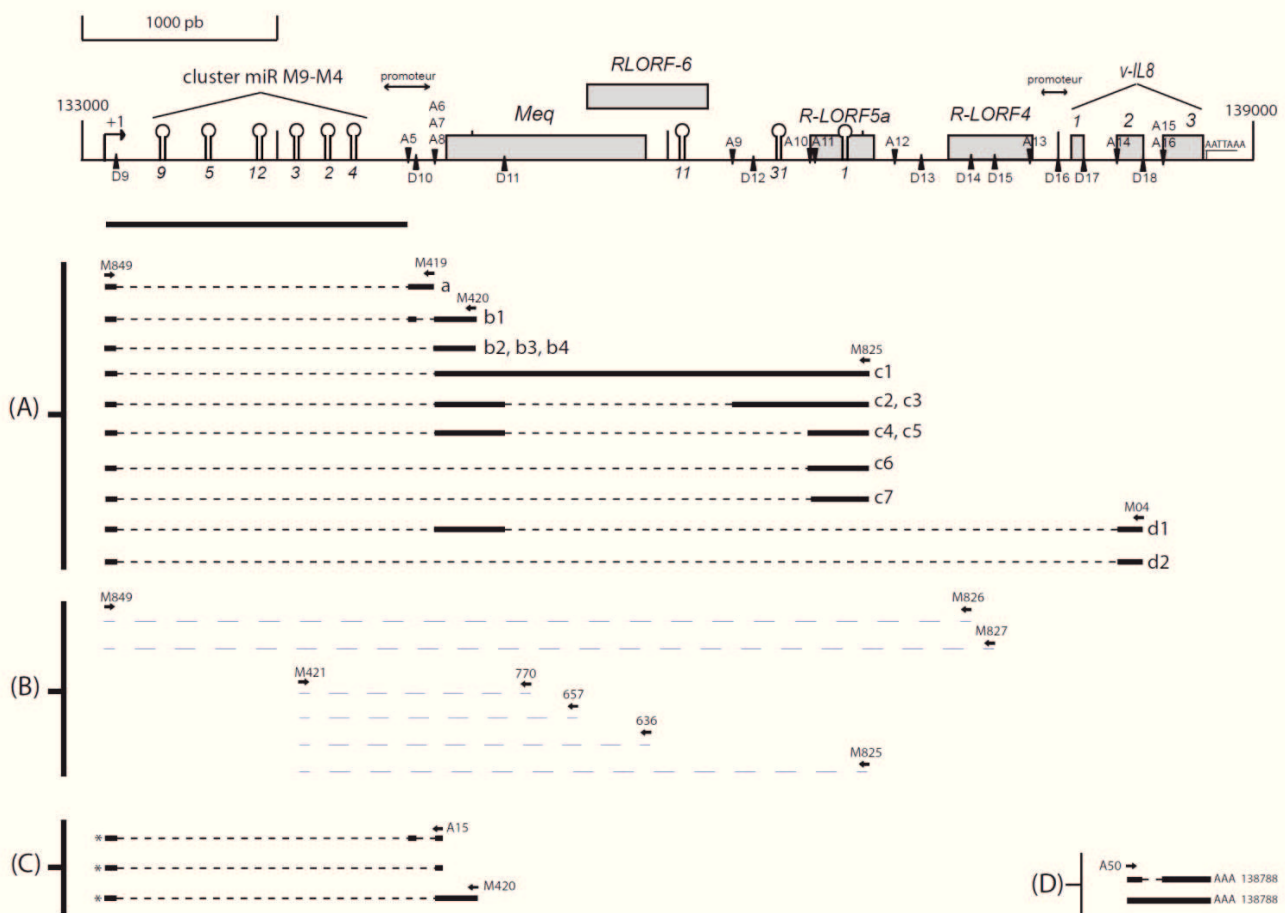


Figure 56: Représentation schématique des transcrits s’initiant en amont de mdv1-miR-M9-M4.

La région comprise entre les bases 133000 et 139000 de GaHV-2 Md5 est représentée dans la partie supérieure du schéma. Les ORF sont indiquées par des rectangles gris et les miARN par des tiges boucles. Les bornes d’épissage accepteurs sont schématisées par des triangles noirs inversés et les bornes d’épissage donneuses par des triangles noirs.

(A) Représentation des fragments obtenus par RT-PCR à l’aide de l’oligonucléotide « sens » M849 et de quatre oligonucléotides « anti-sens » différents (M419, M420, M825 et M04). (B) Représentation des RT-PCR n’ayant pas permis d’amplification spécifique de la région. (C) Représentation des transcrits obtenus par 5’RACE PCR à l’aide des oligonucléotides A15 et M420. (D) Représentation des fragments obtenus par 3’RACE PCR à l’aide de l’oligonucléotide A50, correspondant à l’extrémité 3’ du gène *v-IL8*. Les exons sont représentés par des traits pleins noirs et les introns de type U2 par des pointillés noirs. Les pointillés bleus correspondent aux RT-PCR ne permettant pas d’amplification spécifique de la région. Les amorces utilisées pour chaque amplification sont représentées au dessus du premier transcrit de chaque amplification, l’étoile (*) correspond à l’extrémité 5’ identifiée par 5’RACE PCR et la queue polyadénylée identifiée par 3’RACE PCR est symbolisée par AAA. Seuls les transcrits identifiés à plusieurs reprises au sein d’une même PCR sont représentés. De plus, les transcrits possédant les mêmes exons mais possédant une borne d’épissage alternative proche ne sont figurés que sous la forme d’un seul transcrit.

Amorces PCR*	Transcrit alternatif§	pourcentage†	Taille du transcrit (pb)	Bornes Intron 1	Bornes Intron 2
M849-M419	ARN a	100	198	D9-A5 [#]	
M849-M420	ARN b1	2	316	D9-A5	D10-A8
	ARN b2	2	287	D9-A6	
	ARN b3	20	283	D9-A7	
	ARN b4	76	270	D9-A8	
M849-M825	ARN c1	18	2267	D9-A8	
	ARN c2	18	1077	D9-A7	D11-A9
	ARN c3	36	1090	D9-A8	D11-A9
	ARN c4	9	725	D9-A7	D11-A10
	ARN c5	5	712	D9-A8	D11-A10
	ARN c6	5	371	D9-A10	
	ARN c7	9	407	D9-A11	
M849-M04	ARN d1	10	509	D9-A7	D11-A14
	ARN d2	62	532	D9-A8	D11-A14
	ARN d3	28	191	D9-A14	
M849-M826	--				
M849-M827	--				
M421-770	--				
M421-657	--				
M421-636	--				
M421-M825	--				

Tableau 17: Transcrits identifiés en cellules MSB-1 débutant en amont de mdv1-miR-M9-M4 et permettant d'encoder les miARN du cluster1 de GaHV-2.

* Amorces « sens » et « inverse » utilisées

§ nom des transcrits obtenus pour chaque RT-PCR,

† pourcentage de chaque transcrit identifié au sein d'une même RT-PCR,

nomenclature des bornes donneuses et acceptuses de chaque transcrit identifié.

	exon intron	position IR _L (Md5)
D1*	TATATTTCCT GT AAGGAACG [†]	128506-128525 [‡]
D2 [#]	CATCAAATAA GT AATAAAAT	128562-128581
D3	TGTCGTTTCT GT ATGGTCTC	128754-128773
D4	GGTGGGAACG GT ATGCCATG	129856-129875
D5 [#]	TCAAGTTTGG GT GATCGCAT	130962-130981
D6 [#]	AGAAGTTTTT GT GAGAGGGC	130993-131012
D7 [#]	TATTTTGAAT GT ATGTATAT	132092-132111
D8 [#]	TCGGTTGAAG GT GCGATCTG	132965-132984
D9 [#]	CAAAGTGGG GT AAGGTAAT	133169-133188
D10 [#]	ATACTTTCGG GT CTGTGGGT	134707-134726
D11	GAACACCTAC GT AAGGAAAT	135155-135174
D12 [#]	TTTAGCAGAT GT GAGTTAAT	136422-136441
D13	AAACTAACG GT TCGTGGTT	137289-137308
D14	CTTGTTTTGG GT AATTGGTC	137545-137564
D15	GCGTACTGCA GT AAGTGGTC	137665-137684
D16	GGCATTACGG GT ACGGTTCC	137992-138011
D17	CCTGGAAATG GT AAGTAGGG	138119-138138
D18	CTGAAATTAT GT AAGTAAGC	138430-138449

Tableau 18: Séquence des bornes donneuses de type U2 présentes sur le brin codant les miARN de l'IR_L de GaHV-2.

* nom des bornes D1 à D18 donné en fonction de leur position par rapport au début de l'IR_L (jonction U_L/IR_L),

bornes nouvellement identifiées dans le cadre de notre étude,

† séquence des 20 nucléotides encadrant la borne exon/intron,

‡ position des bornes donnée en fonction du génome de GaHV-2 Md5.

	intron exon	position IR _L (Md5)
A1* [#]	TGCTGCG TAG CTGTTTTGTG [†]	129727-129746 [‡]
A2	ATCGTTG CAG ACCCCGGATG	130891-130910
A3 [#]	GGAAATAT TAG GTAATAGGAA	131734-131753
A4 [#]	TTTCATAC CAG GATCTGGCTG	133052-133071
A5 [#]	AAAATGAC CAG GTGAATTGTG	134661-134680
A6 [#]	TGATTCC TAG GCAGGCGTCT	134797-134816
A7 [#]	TCCTAGG CAG GCGTCTCTTG	134801-134820
A8 [#]	TCTCTTG CAG GTGTATACCA	134814-134833
A9	TTTTTTT AAG AAGGGATATG	136332-136351
A10 [#]	CGATACAC CAG GCACATAGGC	136710-136729
A11 [#]	GAATATAC CAG GCACGGACCT	136746-136765
A12	GTACCGAT TAG TGCCGCGCGT	137153-137172
A13	TTCATTT TAG TGTTTCGATG	137844-137863
A14	AATAATG TAG GCATATCACT	138294-138313
A15	TTTGGTT AAG GTTTTCCCTT	138512-138531
A16	CCTTTTG TAG CTTTGCTCTC	138528-138547

Tableau 19: Séquence des bornes accepteuses de type U2 présentes sur le brin codant les miARN de l'IR_L de GaHV-2.

* nom des bornes A1 à A16 donné en fonction de leur position par rapport au début de l'IR_L (jonction U_L/IR_L),

bornes nouvellement identifiées dans le cadre de notre étude,

† séquence des 20 nucléotides encadrant la borne exon/intron,

‡ position des bornes donnée en fonction du génome de GaHV-2 Md5.

identifiés (Tableau 17). Ils conduiraient à la localisation des miARN mdv1-miR-M9-M4 au sein d'un intron bordé par les bornes D9 et A5 (Figure 56A, tableaux 17, 18 et 19). Toutes ces bornes d'épissage présentent une séquence de type U2 (tableaux 18 et 19), la plus fréquemment rencontrée dans les phénomènes d'épissage. De plus, cette nouvelle borne d'épissage donneuse D9 située en amont du mdv1-miR-M9 semble pouvoir s'épisser avec les quatre bornes accepteuses identifiées en amont de *meq* (A5, A6, A7 et A8). Les deux bornes A7 et A8 correspondant aux transcrits alternatifs ARNb3 et ARNb4 ont été retrouvées majoritairement, dans respectivement 25% et 70% des transcrits (Figure 56, Tableaux 18, 19). Enfin, il faut noter l'existence du transcrit minoritaire ARNb1, qui présentait deux introns dans cette région, le grand intron recouvrant les miARN (bornes D9-A5) et un petit intron bordé par les bornes D10 et A8.

L'utilisation de l'oligonucléotide M825 situé au sein du gène *R-LORF5a* a permis l'identification de sept transcrits alternatifs ARNc1 à c7 de longueur comprise entre 371 et 2267 nucléotides (Figure 56 ; Tableau 17). Le transcrit le plus long ARNc1 (2267pb), représentant 18% des transcrits obtenus avec le couple d'amorce M849-M825 permettrait la transcription du gène *meq* sous la forme d'un grand exon non épissé recouvrant également les trois miARN M11, M31 et M1. Quatre transcrits ARNc2, ARNc3, ARNc4 et ARNc5 (18, 36, 9 et 5 % des clones séquencés) sont composés de trois exons : a) le premier exon se terminant à la borne D9, commun à tous les transcrits des différentes banques, b) un exon intermédiaire recouvrant l'extrémité 5' du cadre de lecture de *meq* (exon-*meq*) compris entre les bornes A7-D11 (ARNc2 et ARNc4) ou A8-D11 (ARNc3 et ARNc5), c) un exon 3' débutant à la borne A9 (ARNc2 et ARNc3) ou A10 (ARNc4 et ARNc5). Les deux transcrits les plus présents (ARNc2 et ARNc3) localiseraient les miARN mdv1-miR-M31 et M1 à l'intérieur d'un exon et le reste des miARN du cluster1 (M9 à M4 et M11) dans deux introns différents, alors que les deux transcrits (ARNc4 et ARNc5) localiseraient uniquement le mdv1-miR-M1 à l'intérieur d'un exon. Les transcrits ARNc6 et ARNc7 correspondent à l'excision d'un grand intron (D9 à A10 ou A11) comprenant tous les miARN du cluster1 à l'exclusion du miR-M1 lequel, à l'instar de la situation observée pour les transcrits ARNc4 et ARNc5 serait contenu dans un petit exon débutant aux bornes A10 (ARNc6) ou A11 (ARNc7).

La recherche de transcrits potentiels comprenant les régions plus en 3' de l'IR_L/TR_L a été effectuée par RT-PCR à l'aide du couple d'amorces M849-M04. Cette dernière a permis l'identification de trois transcrits (ARNd1, ARNd2 et ARNd3) recouvrant 5,5 kpb de la région IR_L et situant l'ensemble des miARN du cluster1 à l'intérieur d'introns. Comme précédemment pour les transcrits ARNc, deux types de transcrits coexistent : deux transcrits ARN d1 et ARNd2 renfermant l'exon-*meq* (bornes A7 ou A8–D11) localisant les miARN dans deux introns distincts, et un transcrit ARNd3 résultant de l'excision d'un grand intron (bornes D9-A14) renfermant la totalité des miARN du cluster1.

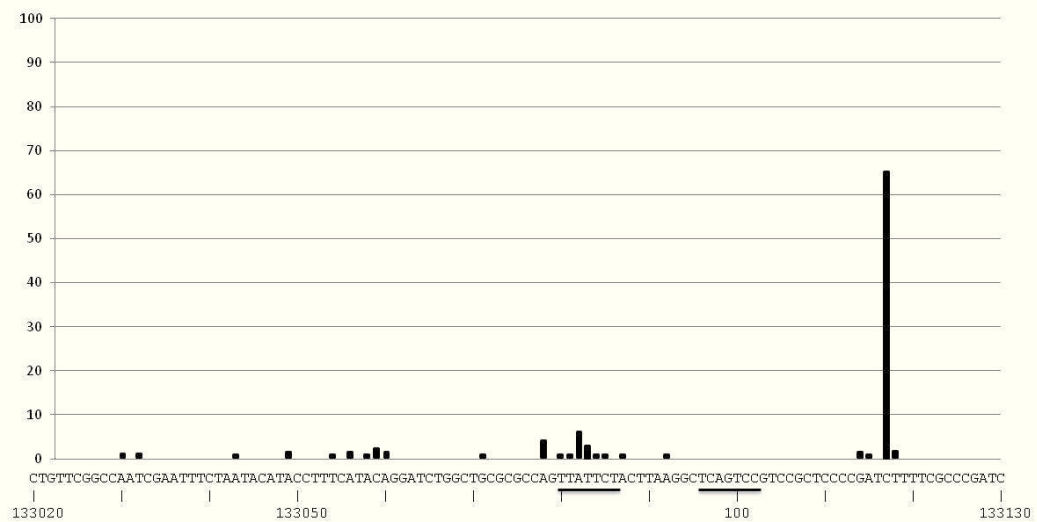


Figure 57: Représentation schématique de l'ensemble des TSS identifiés par 5'RACE PCR dans la région p-miR-M9-M4 en cellules MSB-1.

Le pourcentage des TSS obtenus lors des 5'RACE PCR réalisées avec les amorces M416, A15 et M420 sont représentés par des barres au dessus de chaque base correspondant au moins à une initiation. La séquence comprise entre les bases 133020 et 133130 du génome de GaHV-2 Md5 est représentée en dessous. Les INR potentielles (YYANWYY) présents sur la séquence sont soulignés.

Enfin, les RT-PCR réalisées avec deux oligonucléotides localisés dans le gène *R-LORF4* se sont révélées infructueuses, semblant indiquer que *R-LORF4* n'est pas détectable sous la forme d'un exon dans les transcrits obtenus avec les couples d'amorces M849-M826 et M849-M827 (Tableau 17 ; figure 56B). De même, aucun amplicon n'a été détecté en utilisant une amorce « sens » localisée entre mdv1-miR-M3 et M2 ainsi que des oligonucléotides « anti-sens » situés dans l'intron en aval de l'exon-*meq* (Tableau 17; figure 56B). Ainsi, ces résultats sembleraient faire ressortir l'absence d'épissage alternatif en aval de mdv1-miR-M3 et surtout souligneraient l'importance de la borne donneuse D9 lors des mécanismes d'épissage différentiels dans cette région (Tableau 17 ; figure 56B).

Pour confirmer les résultats précédents, afin de déterminer l'initiation de ces transcrits, deux 5'RACE PCR ont été réalisées à l'aide d'oligonucléotides localisés au sein du promoteur de *meq* (A15), ou au sein de la région codante de *meq* (M420) (Figure 56C). Toutes ces 5'RACE PCR réalisées à partir d'ARN totaux extraits de cellules MSB-1 s'initiaient en amont du miARN mdv1-miR-M9 permettant de valider la présence d'une région promotrice en amont des miARN. Cependant, en plus de l'initiation majoritaire (position 133117 du génome de GaHV-2 Md5), il a été observé différentes initiations minoritaires en amont de l'initiation dominante (Figure 57). En effet, 65% des transcrits s'initiaient en position 133117 mais des initiations dispersées ne représentant pas plus de 7% par TSS et s'étendant sur environ 90pb ont également été retrouvées (Figure 57). Enfin, il faut noter que 15% de ces TSS sont situés au niveau d'une INR (133080-YYANWYY-133086) (Figure 57). Pour terminer, les profils d'épissage obtenus lors de ces RACE-PCR ont parfaitement confirmé les événements décrits lors des RT-PCR ciblées, c'est à dire l'importance de la borne D9 et l'existence de transcrits alternatifs mettant également en jeu les bornes A5, A6, A7, A8 et D10 (Figure 56C).

Pour confirmer dans notre modèle les résultats décrits dans la littérature, une 3'RACE PCR sur le gène *vIL-8* a été réalisée en cellules MSB-1 afin d'identifier l'extrémité 3' des transcrits de *vIL-8* et potentiellement des transcrits s'épissant avec l'exon2 de *vIL-8* (Figure 56D). Pour cela, l'utilisation de l'oligonucléotide « sens » A50 situé au début de l'exon2 a permis de retrouver l'extrémité 3' de *vIL-8* localisée en position 138788 du génome de GaHV-2 Md5 soit à 31 pb en aval de la fin du signal de polyadénylation AATTAAA (Figure 56D), précédemment décrite par Parcells et ses collaborateurs (Parcells *et al.*, 2001).

2.1.3. Recherche de longs transcrits recouvrant la région IR_L/TR_L de GaHV-2.

Après l'identification de transcrits permettant l'expression des miARN du cluster 1 sous la forme d'un exon ou d'introns et pouvant recouvrir la région IR_L de mdv1-miR-M9 jusqu'à *vIL-8*, j'ai recherché, si parmi les transcrits s'initiant plus en amont, certains pouvaient permettre également la transcription du cluster1 des miARN.

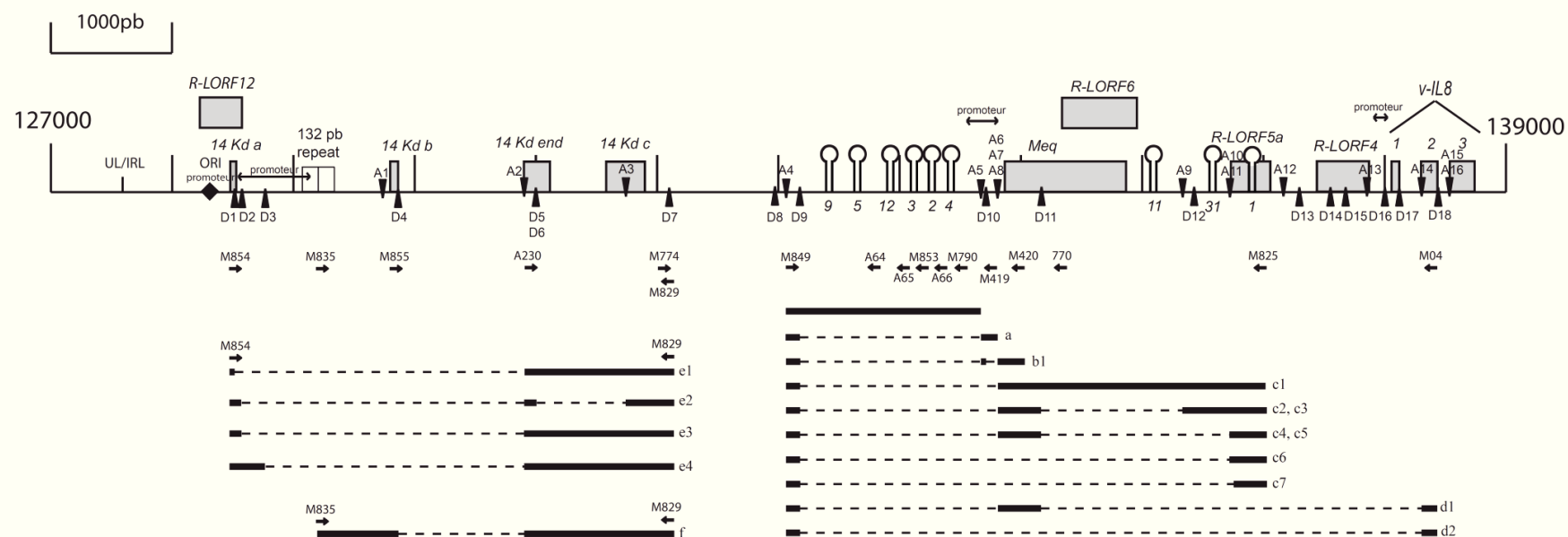


Figure 58 : Recherche des transcrits spécifiques des gènes des protéines 14kD.

La région comprise entre les bases 127000 et 139000 de GaHV-2 Md5 est représentée dans la partie supérieure, les gènes sont représentés par des rectangles gris et les miARN par des tiges boucles. Les bornes d'épissage accepteuses sont schématisées par des triangles noirs inversés et les bornes d'épissage donneuses par des triangles noirs. Les transcrits précédemment identifiés dans le cadre de notre étude sont figurés à droite.

Les transcrits ont été amplifiés à l'aide des amorces « sens » M854 et M855 et de l'amorce « inverse » M820. Les exons sont représentés par des traits pleins noirs et les introns de type U2 par des pointillés noirs. Seuls les transcrits identifiés à plusieurs reprises au sein d'une même PCR sont représentés. De plus, les transcrits possédant les mêmes exons mais avec une borne d'épissage alternative proche ne sont figurés que sous la forme d'un seul transcrit.

En préambule à la recherche de longs transcrits couvrant la totalité de la région IR_L/TR_L, j'ai vérifié la présence des transcrits codant les gènes de la famille des 14kD précédemment décrits par Hong et ses collaborateurs (Hong & Coussens, 1994). Pour cela, deux RT-PCR ont été réalisées à l'aide des oligonucléotides « sens » M854 et M855 localisées respectivement au sein des exons *14kDa* et *14kDb* et de l'oligonucléotide « anti-sens » M829 situé en aval de *14kDc* (Figure 58). En plus des trois transcrits ARNe1, ARNe4 et ARNf identifiés auparavant (Hong & Coussens, 1994), deux nouveaux transcrits (ARNe2 et ARNe3) ont été observés. Ils présentaient une nouvelle borne donneuse D2 au sein de l'intron entre *14kDa* et *14kDend* conduisant à un exon1 différent des deux décrits dans la littérature (Figure 58), gommant le codon stop de l'ORF R-LORF12 mais sans le mettre en phase avec *14kDend*. De plus, le transcrit ARNe2 présentait également un second intron délimité par les bornes D6-A3 localisées au niveau des gènes *14kDend* et *14kDc*.

Suite à l'identification de nouveaux transcrits alternatifs pour la famille des gènes 14kD et sachant qu'un long transcrit de 11kb non codant (LAT) avait été identifié au sein des régions IR_S/TR_S, j'ai recherché l'existence potentielle de longs transcrits couvrant la région TR_L/IR_L par RT-PCR à l'aide d'amorces localisées dans le début de *14 kDa* (M854) et dans le second exon de *vIL-8* (M04). Ainsi, neuf longs transcrits fortement épissés ARNg1 à ARNg9 recouvrant une région longue d'environ 10kpb ont été mis en évidence (Figure 59A). Parmi ces transcrits, les ARNg1 et ARNg2 présentaient un épissage classique de type U2 avec cinq exons potentiels localisés au niveau des différentes ORF déjà décrites, l'exon *14kDa*, l'exon *14kDb*, l'exon *14kDend*, ainsi que l'exon-*meq* et l'exon 2 de *vIL-8* (Figure 59A). Il est intéressant de noter la présence d'un épissage entre l'exon *14kDa* et l'exon *14kDb* au sein d'un de ces transcrits, l'ARNg1 (Figure 59A), qui contraste fortement avec les résultats de la littérature pour lesquels *14kDb* n'était transcrit que lors d'une initiation propre mais jamais en association avec *14kDa* (Hong & Coussens, 1994). Par ailleurs, 14 des 16 transcrits séquencés présentaient des bornes non canoniques pour un épissage de type U2. En effet, les bornes de ces introns présentaient une séquence de type AT/AC, spécifique des introns de type U12 au lieu de bornes GT/AG spécifiques des introns de type U2 (Tableau 20). De plus, ces transcrits présentaient un profil différent des grands transcrits présentés précédemment avec des bornes accepteuses et donneuses relativement variables (Figure 59A). Effectivement, six bornes donneuses potentielles de type U12 ont été identifiées en aval de *14kDa* (d1-d6 ; Tableau 20) et quatre bornes accepteuses potentielles de type U12 (a3 à a6) ont été retrouvées au sein de l'exon 2 de *vIL-8* (Figure 59A). Il est intéressant de remarquer pour les ARNg3, ARNg4 et ARNg5, la présence d'une borne accepteuse a1 (U12) au sein même de l'exon 1 de *vIL-8* suivi d'un épissage classique de type U2 entre l'exon 1 et l'exon 2 de *vIL-8*, ce qui n'avait jamais été identifié dans la littérature ou dans les transcrits précédemment décrits dans ce manuscrit (Figure 59A). Pour finir, l'ensemble de ces transcrits fortement épissés, par un mécanisme de type U2 ou U12, permettrait la transcription de tous les miARN de l'IR_L au sein d'un grand intron unique ou de deux introns quand l'exon-*meq* est présent.

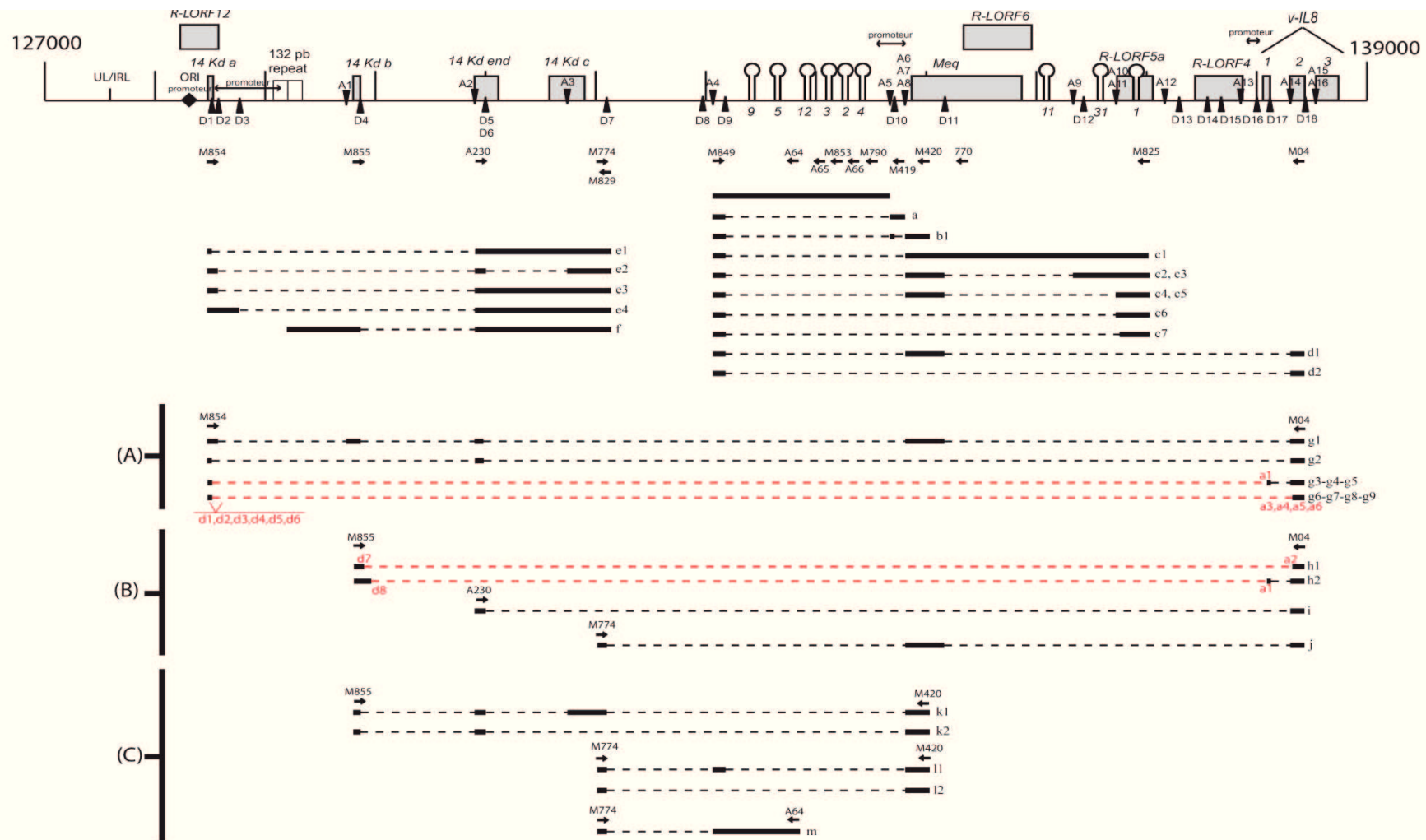


Figure 59: Représentation schématique de l'ensemble des transcrits encodant les miARN du cluster 1 de GaHV-2 en cellule MSB-1.

Les transcrits obtenus dans le cadre de mon étude ainsi que la région comprise entre les bases 127000 et 139000 de GaHV-2 Md5 sont schématisés en haut de la figure. Les ORF sont schématisées par des rectangles gris et les miARN par des tiges boucles. Les bornes d'épissage accepteurs de type U2 sont schématisées par des triangles noirs inversés et les bornes d'épissage donneuses de type U2 par des triangles noirs. Les exons sont représentés par des traits pleins noirs, les introns de type U2 par des pointillés noirs et les introns de type U12 par des pointillés rouges. Les amorces utilisées pour les différentes amplifications sont représentées au dessus des transcrits. (A) Représentation des transcrits obtenus par RT-PCR à l'aide du couple d'amorces M854/M04. (B) Représentation des transcrits obtenus par RT-PCR à l'aide des l'oligonucléotides « sens » M855, A230 et M774 et de l'oligonucléotide « anti-sens » M04. Les bornes d'épissages des introns de type U12 (d1 à d8 et a1 à a6) sont directement représentées sur les transcrits présentant un épissage de type U12. (C) Représentation des transcrits obtenus par RT-PCR à l'aide des amorces « sens » M855 et M774 et « inverse » A64 et M420. Seuls les transcrits identifiés à plusieurs reprises au sein d'une même PCR sont représentés. De plus, les transcrits possédant les mêmes exons mais possédant une borne d'épissage alternative proche ne sont figurés que sous la forme d'un seul transcrit.

séquences*	bornes donneuses [†]	bornes accepteuses [‡]
g3	d1	a1
g4	d2	a1
g6	d3	a5
g7	d4	a3
g8	d4	a6
g5	d5	a1
g9	d6	a3
h1	d7	a2
h2	d8	a1

	exon intron	position IR _L (Md5)
d1 [†]	AATTACTAGC AT TCGATAAG [°]	128579-128598 [§]
d2	ACTAGCATTC GATAAGCAAA	128583-128602
d3	TTTTATTTAC GGCGATATTT	128638-128657
d4	TTCGTCGGTC AGTCCCGACG	128672-128691
d5	GACCGTCAAC ATCGACCCGG	128707-128726
d6	GGGCAGACCC ATGGGTGTCTG	128739-128758
d7	ACGGTATGCC ATGACAGATT	129863-129882
d8	TATCGTGATC ATACTCAGGT	130226-130245

	intron exon	position IR _L (Md5)
a1 [‡]	TTCATAGT AC AGATCTATTT [°]	138096-138115 [§]
a2	CGCTGTCTG AC AAGAGGTGCA	138324-138343
a3	CAAGAGGT GC AAGTGCCTGA	138333-138352
a4	ATCGGCCT AC TGGCTTGGGG	138362-138381
a5	GGGGCCTAT TT ATCGCTGTTG	138379-138398

Tableau 20: Séquences nucléotidiques des bornes d'épissage des neuf introns potentiellement de type U12 présent dans l'IR_L de GaHV-2.

* Séquences présentant un intron potentiel de type U12,

^{†,‡} nom des bornes d1 à d8 ([†]) et a1 à a5 ([‡]) donné en fonction de leur position par rapport au début de l'IR_L (jonction U_L/IR_L),

[°] séquence des 20 nucléotides encadrant la borne exon/intron

[§] position des bornes donnée en fonction du génome de GaHV-2 md5.

Afin de conforter les résultats précédents, une série de RT-PCR a été réalisée en utilisant les oligonucléotides « sens » localisés à l'intérieur des ORF de *14kDb* (M855) et de *14kDend* (A230) ou en aval de *14kDc* (M774) et l'oligonucléotide « anti-sens » M04 situé dans l'exon 2 de *vIL-8* (Figure 59B). A l'instar des transcrits de type U12 observés entre *14kDa* et *vIL-8*, les deux types de transcrits ARNh1 et ARNh2 obtenus avec le couple M855/M04 présentaient également des introns potentiellement de type U12 avec une séquence de type AT/AC. De plus, tout comme les ARNg3 à g9, ces deux transcrits présentaient des événements d'épissage alternatif avec les exons 1 ou 2 de *vIL-8* (Figure 59B). A l'inverse, l'utilisation des oligonucléotides A230 et M774 a permis l'amplification de transcrits de type ARNi et ARNj. Les événements d'épissage correspondants de type U2, raccordaient les exons spécifiques *14kDend* et *14kDc* à l'exon 2 de *vIL-8* soit directement pour l'ARNi, soit par l'intermédiaire de l'exon-*meq* pour l'ARNj (Figure 59B).

En parallèle de l'identification des longs transcrits s'épissant avec l'exon 2 de *vIL-8*, nous avons recherché des transcrits s'épissant avec le gène *meq*. Ainsi, à l'aide de l'oligonucléotide « anti-sens » M420 et des oligonucléotides « sens » M855 et M774, quatre transcrits ARNk1, ARNk2, ARNi1 et ARNi2 ont été identifiés. A l'exception de l'ARNi1 qui présentait un petit exon situé juste en amont des miARN et possédant les bornes A4-D9, ces transcrits ont permis de confirmer les exons précédemment décrits dans ce manuscrit et la transcription du gène *meq* par épissage à partir des gènes *14kD* que nous avons mise en évidence (Figure 59C).

Enfin, comme les différents transcrits de longue taille (Figure 59B et C) s'initiant au niveau des gènes 14kD présentaient des profils se recoupant totalement avec ceux des transcrits plus spécifiques de la région 3' (ARNa à ARNd), j'ai recherché l'existence d'un exon potentiel encodant les miARN mdv1-miR-M9 à M4. Ainsi, des RT-PCR ont été réalisées avec l'oligonucléotide « sens » M774 ainsi que les oligonucléotides A64, A65, M853, A66 et M790 situés entre les différents miARN en aval de mdv1-miR-M5. L'utilisation de l'oligonucléotide « anti-sens » A64 situé entre mdv1-pré-miR-M5 et M12 a permis l'amplification du transcrit ARNm de 867pb contenant les deux miARN mdv1-miR-M9 et M5 au sein d'un exon (Figure 59C). Ainsi, l'utilisation de la borne accepteuse A4 retrouvée fonctionnelle au sein du transcrit ARNi1 semblerait pouvoir permettre la formation d'un exon contenant tout ou une partie des miARN mdv1-miR-M9 à M4. Des essais complémentaires visant à augmenter le nombre de séquences des transcrits de cette région sont en cours pour évaluer complètement cette hypothèse. En effet, tous les transcrits obtenus avec les différentes amorces situées en aval du mdv1-pré-miR-M5 correspondaient à des amplifications de transcrits du brin complémentaire du génome viral. De plus, ce nombre était trop faible pour pouvoir conclure à une absence réelle de transcrit couvrant la totalité des mdv1-pré-miR-M9 à M4.

Pour conclure, si l'on considère uniquement les transcrits épissés, excepté la formation potentielle d'un exon contenant les miARN mdv1-miR-M9 à M4, qui reste à être confirmée, les miARN du cluster1 sont constamment présents au sein d'introns selon deux cas de figure. Tout d'abord, un épissage entre les bornes D5, D6, D7 ou D9 et les bornes accepteuses présentes en amont

de *meq* (A7-A8) répartit les miARN du cluster1 en deux introns distincts. Ensuite, un second profil, présentant un intron délimité par les bornes donneuses D5, D6 et D7 et la borne accepteuse A14 localise l'ensemble des miARN au sein d'un intron unique (Figure 59A).

2.2. Etude comparative de transcrits encodant les miARN de la région IR_L/TR_L entre les phases de latence et de réplication de GaHV-2.

Afin de comparer les initiations et les profils des transcrits précédemment identifiés à partir des cellules MSB-1 représentatives de la phase de latence, avec ceux observés lors d'une infection lytique de GaHV-2, trois 5'RACE PCR ont été initiées au sein des gènes *meq* et *vIL-8* à partir des ARN totaux extraits de CEF infectés par la souche GaHV-2 RB-1B.

Tout d'abord, comme précédemment pour les cellules MSB-1, deux 5'RACE PCR ont été amorcées à partir du gène *meq* à l'aide de deux oligonucléotides situés en amont de celui-ci (A15) ou dans sa région codante (M420). A l'instar des données observées pour les cellules MSB-1, le séquençage de plus de 40 clones pour chacune des deux RACE-PCR a fait ressortir majoritairement une initiation des messagers au niveau du TSS-133117. En effet, ce TSS représentait 63% et 75% des clones obtenus aux RACE-PCR amorcées respectivement par A15 et M420, le reste des clones présentant des extrémités potentielles pour les ARNm réparties aléatoirement sur la région précédemment décrite comme promotrice pour le gène *meq*. Aussi, comme ces extrémités présentaient des localisations différentes, elles ne semblaient pas correspondre à une région d'initiation spécifique, mais plutôt au bruit de fond de la technique.

Sachant que le gène *vIL-8* est essentiellement exprimé durant l'infection lytique à partir d'un promoteur spécifique s'initiant en position 138049 de GaHV-2 Md5, localisé entre *R-LORF4* et *vIL-8* (Parcells *et al.*, 2001), et afin de tester l'utilisation potentielle de promoteurs alternatifs lors du cycle lytique ou de la latence, j'ai réalisé des RACE PCR avec l'amorce M04, situé dans l'exon 2 de *vIL-8*, à partir des ARN totaux extraits de CEF infectés par la souche GaHV-2 RB-1B ou de cellules MSB-1 (Figure 60).

Les 5'RACE PCR réalisées à partir des ARN extraits de cellules MSB-1 ont montré qu'une majorité des transcrits (78%) s'initiaient en amont de mdv1-miR-M9 principalement au niveau du TSS-133117 (60%), avec quelques initiations dispersées plus en amont à l'instar des résultats décrits précédemment pour les cellules MSB-1 avec les oligonucléotides A15 et M420 (Figure 57). Les 22% restants s'initieraient au sein du promoteur de *vIL-8* au niveau du TSS déjà décrit par Parcells et ses collaborateurs (Parcells *et al.*, 2001), localisé en position 138049 du génome de GaHV-2 Md5 (Figure 60).

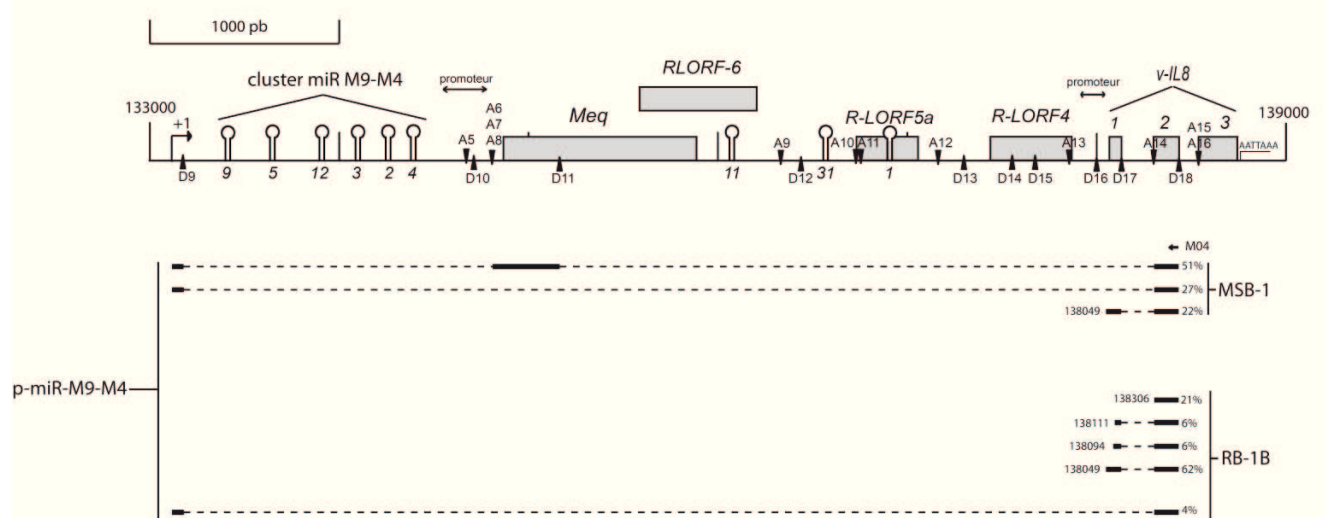


Figure 60: Comparaison des extrémités 5' des transcrits encodant l'exon 2 de *v-IL8* obtenues par 5' RACE PCR durant la phase de latence ou la réplication de GaHV-2.

La région comprise entre les bases 133000 et 139000 de GaHV-2 Md5 est représentée dans la partie supérieure, les ORF sont représentées par des rectangles gris et les miARN par des tiges boucles. Les bornes d'épissage accepteurs sont schématisées par des triangles noirs inversés et les bornes d'épissage donneuses par des triangles noirs. L'ensemble des transcrits ont été obtenus par 5'RACE PCR à l'aide de l'oligonucléotide M04 à partir d'ARNm extraits de cellules MSB-1 (phase de latence) ou de CEF infectés par la souche GaHV-2 RB-1B (phase de réplication). Les exons sont représentés par des traits pleins noirs et les introns de type U2 par des pointillés noir. Le pourcentage de chaque transcrit est représenté à droite de celui-ci.

Par contre, les proportions des messagers s'initiant à partir du promoteur *vIL-8* (TSS-138049) et du promoteur miARN (TSS-133117) étaient complètement inversées dans les banques correspondant aux ARNm de CEF infectées. Effectivement, lors du cycle réplcatif de GaHV-2 en cellules CEF, seuls 4% des transcrits s'initiaient en amont de mdv1-miR-M9 au niveau du TSS-133117, tandis que 74% s'initiaient au niveau du promoteur de *vIL-8* (Figure 60). Plus précisément, 62% des extrémités 5' des transcrits correspondaient au TSS-138049 et 12% étaient localisées à proximité au niveau des nucléotides 138111 et 138094. Par ailleurs, il faut noter que nous avons identifié une initiation pour 21% des messagers directement au niveau de l'exon2 de *vIL-8* en position 138306 du génome de GaHV-2 Md5 (Figure 60).

2.3. Etude de la région promotrice des miARN de la région IR_L de GaHV-2.

2.3.1. Analyse des éléments *cis*-régulateurs potentiels de la région promotrice p-miR-M9-M4.

Les éléments *cis*-régulateurs présents au sein de la région promotrice p-miR-M9-M4 ont été recensés par analyse bioinformatique à l'aide du logiciel Genomatix (www.genomatix.de/). La région promotrice définie arbitrairement du nucléotide 132013 au nucléotide 133627 soit 1104 pb en amont du TSS majoritaire (position 133117 sur génome GaHV2-Md5) et 509 pb en aval correspond à l'intégralité de la région non codante située entre le gène *l4kDc* et mdv1-miR-M5. Ainsi, 80 sites *cis*-régulateurs correspondant à des éléments de réponse (ER) ubiquitaires impliqués dans la transcription basale ainsi que des sites plutôt spécifiques des cellules lymphocytaires ont été identifiés (Figure 61 ; Tableau 21). Parmi les éléments ubiquitaires, sept boîtes TATA, situées en position -984, -522, -490, -210, -25, +242 et +300, et deux boîtes CAAT/NFY localisées en positions -85 et +416 ont été identifiées ainsi que 14 séquences permettant la fixation de facteurs de la famille E2F réparties entre -600 et +62 et quatre boîtes GC permettant celle du facteur Sp1 en position -708, -702, -203 et +379. Il est intéressant de remarquer la présence d'une boîte TATA localisée 25 pb en amont de l'initiation majoritaire 133117 ainsi qu'une boîte CCAAT 85 nucléotides en amont de cette dernière. En effet, ces éléments *cis*-régulateurs dont la position par rapport au TSS est en faveur d'une potentielle fonctionnalité font partie des éléments classiques retrouvés lors d'une initiation unique de la transcription par l'ARN polymérase II.

En plus de ces éléments, nous avons dénombré des sites de liaison à divers facteurs particulièrement importants pour l'expression des gènes dans les cellules sanguines et les lymphocytes

132013 AACGATTGCGGAAATACGGCGGTCGCCGAGACGAACGCGGACGTGTAGCGGGGACGTTTCATTGTCTTTGT
 --> →ETSF (SPI1/Pu1)
 cMyb01←
 132083 TTTTCTAATTATTTTGAATGTATGTATATTTTCAGCCTCATCCTGTAAATCGGTGAGCATTAAAGGT
 →OCT1 →OCT1pou →TATA
 OCT1← ETSF (PDEF)←
 132153 ACGGATATTGGTTCACCTGTATGCGTGTTCATTTTATCTTTTCGTATACTCCCTTCGCGTACGGTACGATC
 GATA1←
 132223 AAGGTTAAAAAAGTTACTGCAATGTACAGTAACCGCCGTAAATTAACCTGCGGTGTCGGTTCAT
 →cMyb02 cMyb02←
 132293 GCGTCTCGATATATGTTCTGTGCGCTTCTCACTAGCTGGCGGCATCTACAATTTTGTCTGTTTGTCTGT
 →Trq-1 ETSF (FLI)←
 132363 TTGCCACTAACTCAATGATTTCCTGGACAGTAGCGCGGGGGAGGGGGTGTGTTATATTTTCGAGAGCGT
 →SPI (GC)
 ETSF (ELK1)← →SPI (TIEG)
 132433 TCGGCCGGGTTTGATCCTGGAATTATGAAACCGTGGCGACTGCGTTTTTATTTTGTGTTGTTATTCTA
 →STAT1 →CEBP
 STAT3←
 132503 TACGTTCCAATGGCGCGAAATGGGGTAGTTTTTTTTCGGGCAGCTCCTCAGGATTATACGGAATTGTT
 →E2FF →E2FF
 E2FF← E2FF←
 132573 TATAACGTCATTGGGTATTGAAGGTATTTCCCTTTATATAAGTTTTTTTAAAGTCGCTGTTCCCATGA
 →vMyb →TATA →NFkB →TATA→Trq-2
 NFAT ; ETSF (SPIB)← Ikaros←
 132643 GAAATTTTGTAACTGATATTGCTTACTCATAGCTGATGCCCGGAATTTTAGCTTAAGTGAATCACCTT
 →GATA1 →AP1 →STAT1 ; ETSF (ELK1)→AP1
 →vMyb AP1← STAT3←
 132713 ATAGTGCAGTGCCTTTGTATGTCTGGGGCCCTGATAGACAGATGCCCTTTCTTAAGTGTGATTTATCATC
 →GATA →Trq-3 →OCT1
 GATA1←
 132783 TCATTACCCCTCGCCTCAGCATAAATATGCGATCGGTGTGACGATTACTCAGAACCCTCTTCCTAATTA
 →OCT1
 OCT3← ETSF (SPIB)← OCT1←
 132853 GCAGAACCCAGCAATTATACTTTAGAGGCCTTTATGTGAGAGGCCCTGTAAATTTTACCCGCCCGAAC
 →Trq-4 →TATA
 E2F← SPI←
 132923 ACTTTCCAACTAACGTGTCGATTGTGTCATCTAGAGACTCTTATCGGTTGAAGGTGCGATCTGTGGGCTTGA
 →cMyb
 GATA← cMyb←
 132993 TTACGCATAATCGTGCGCGAATTGCCCTGTTCGGCCAATCGAATTCTAATACATACCTTTCATACAGGA
 →OCT1pou →E2FF →caat/NFY
 133063 TCTGGCTGCGCGCCAGTTATTCTACTTAAGGCTCAGTCCGCTCCGCTCCCGATCTTTTCGCCCGATCCCT
 →E2FF →TATA →E2FF
 E2F← E2FF←
 cMyb←
 133133 CTACCAGCCGGCTTTTCGATCGTGACATTTAGAACCAAGTGCGGGTAAGGTAATCATTCGCCGATAACT
 →E2FF →GATA1
 →STAT3
 STAT1←
 133203 AAGAGACCTGATCTCCGGGAGTGTAAACATATCGGGTAGATCTAATGAAGTGGGCCAAACAAAATGGGAAA
 →GATA3 →E2F
 OCT1←
 133273 GTAGATTTTAACTTTTCTAGTCTACTCATATACATTTCCGTACGCGGGCGATGCACGAGATAATTCTC
 →GATA3 →GATA1
 OCT1←
 133343 ATTCCGATATATAATTTGTGCCGAGCGGTCTTTCTCCTTCCGCCCGGAGTTCACTGTATCGTACGTGT
 →TATA →Trq-5
 OCT2← ETSF (SPI1/Pu1)←
 133413 AAACCTCCGAGGGCAGGAAAAAGTGCCCCGACTCACATACTTCAAAATATAATTTTGAAGGC TAATGAGC
 →E2F
 133483 ATCGTGGGGGAGGGGGTGTATGTGTGAGCAGTGGTGTAAACCAATATTAGGAGAACGTGAAAAAA
 →SPI (GC) →caat/NFY
 cMyb← →E2F
 133553 AAAAGAGGGGGGAGACGGAAGCTTTTAAAGTTTATCAACTGTTCATGAACCGTATGCGATCACAT
 →cMyb
 GATA1←
 133623 TGAC

Figure 61 : Localisation des différents éléments *cis*-régulateurs présents sur la séquence de la région promotrice p-miR-M9-M4.

Les nucléotides soulignés représentent la séquence reconnue par les facteurs de transcription. Les nucléotides en caractère gras figurent les points d'interaction entre l'élément de réponse et le facteur de transcription associé. Le nucléotide surligné en rouge correspond au site d'initiation majoritaire (133117). Les séquences TAATGARAT like sont surlignées en jaune et les microARN sont surlignés en vert. Le début des promoteurs Trq-1, Trq-2, Trq-3, Trq-4 et Trq-5 sont représentés par des nucléotides surlignés en gris. Abréviations : ETSF : Ets-factor ; OCT : Octamer Binding factor ; SPI : Stimulated protein 1 ; STAT : Signal transducer and activator of transcription ; CEBP : CAAT enhancer binding protein ; NFAT : Nuclear factor of activated T-cells ; NFkB : Nuclear factor Kappa B ; NFY : Nuclear factor Y.

Sites	TATA*	CAAT/NFY*	E2FF*	Spl *	Ikaros [#]	STAT1 [#]	STAT3 [#]	AP1 [#]
nombre	7	2	14	4	1	3	3	5
localisation	-984 [§] , -522, -490, -210, -25, +242 +300	-85, +416	-600, -597, -575, -572, -204, -105, -43, -42, +7, +8, +62, +141, +312, +430	-203, -708, -702, +379	-479	-664, -429, +78	-666, -436, +80	-445, -444, -410, -340, -285
Sites	GATA [#]	Oct-1 [#]	c-Myb [#]	v-Myb [#]	NFκB [#]	NFAT [#]	CEBPβ [#]	ETSF [#]
nombre	10	10	8	2	1	1	1	8
localisation	-928, -461, -369, -338, -152, +83, +127, +163, +220, +477	-1026, -1025, -1007, -330, -316, -266, -265, -115, +132, +284	-1083, -857, -833, -178, -149, -38, +405, +483	-536, -457	-517	-516	-857	-1092, -992, -842, -731, -427, -271, +155, +267

Tableau 21 : Localisation des différents sites cis-régulateurs de la région promotrice p-miR-M9-M4.

* Eléments de réponse ubiquitaires,

[#] éléments de réponse spécifiques des cellules sanguines et lymphocytaires.

[§] localisation des sites par rapport au TSS majoritaire (TSS-133117),

promoteur	amorce sens [†]	amorce anti-sens [‡]	Taille (pb)	Positions *
p-miR-M9-M4	M415	M431	1615	132013-133627
Trq-1	M442	M431	1321	132307-133627
Trq-2	M443	M431	1002	132626-133627
Trq-3	A78	M431	860	132768-133627
Trq-4	M444	M431	724	132886-133627
Trq-5	M445	M431	240	133388-133627
p-court-1	M415	M542	1395	132013-133407
p-court-2	M415	M476	687	132013-132699
Trq-3-court	A78	M542	640	132768-133407
Trq-4-court	M444	M542	522	132886-133407

Tableau 22 : Tableau récapitulatif des différents promoteurs réalisés à partir de la région promotrice p-miR-M9-M4.

[†] Amorce « sens » ayant permis l'amplification du promoteur,

[‡] Amorce « anti-sens » ayant permis l'amplification du promoteur,

* position des promoteurs donnée à partir de l'IR_L de GaHV-2 Md5.

(Tableau 21). Ces sites sont intéressants soit pour leur capacité à se lier à des facteurs clefs de la croissance et de la prolifération comme (STAT (signal transducer and activator of transcription) et c-Myb) soit pour leur implication dans le développement et la différenciation des lymphocytes et des cellules sanguines comme Ikaros, GATA, Oct1 et c-Myb. Deux sites potentiels de fixation de v-Myb ainsi qu'un site fixant NFκB et un site CEBPβ (CAAT Enhancer Binding Protein beta) sont également retrouvés. De plus, certains des sites identifiés pourraient permettre la fixation potentielle de transactivateurs viraux tels que la protéine Meq sur les sites AP1 ou la protéine VP16 sur les sites TAATGARAT like (Figure 61).

2.3.2. Détermination de la région promotrice minimale de p-miR-M9-M4.

2.3.2.1. Construction des formes tronquées du promoteur p-miR-M9-M4.

Afin de définir le « core » promoteur, soit la région minimale permettant une activité promotrice maximale, la région promotrice p-miR-M9-M4 a été amplifiée par PCR à l'aide des oligonucléotides « sens » M416 et « anti-sens » M431 à partir d'ADN génomique de la souche virale GaHV-2 RB-1B. Elle a ensuite été insérée dans le vecteur pGL3 basic, en amont du gène rapporteur luciférase firefly, pour former le vecteur pGL3-p-miR-M9-M4. Puis, différentes troncatures ont été générées à partir des extrémités 5' et 3' de la construction pGL3-p-miR-M9-M4 par PCR à l'aide d'une batterie d'oligonucléotides (Tableau 22). Ainsi les amplicons correspondant aux formes Trq-1, Trq-2, Trq-3, Trq-4, Trq-5 (Figure 62 ; Tableau 22), ont été obtenus avec l'oligonucléotide « sens » M415 et respectivement les oligonucléotides « anti-sens » M442, M443, A78, M444 et M445. Ainsi, cette approche permettait non seulement de définir le « core » promoteur tout en testant globalement par exclusion la fonctionnalité potentielle des différents ER. A titre d'exemple, la forme tronquée 1 de 1321pb (Trq-1 : -811 → +509) excluait trois sites c-Myb, trois sites Oct-1 et une boîte TATA et la forme tronquée 2 de 1002 pb (Trq-2 : -492 → +509) permettait de tester en plus des ER de la forme Trq-1 une autre boîte TATA, une boîte GC (SP1), deux ER pour STAT, un ER pour CEBPβ et un pour v-Myb. Sur le même principe, la forme tronquée 3 de 860pb (Trq-3 : -351 → +509) permettait d'exclure trois sites AP-1, deux sites STAT et un site v-Myb et la forme tronquée 4 de 742 pb (Trq-4 : -232 → +509) trois ER de la famille des Oct et trois ER pour AP-1. Enfin, la forme tronquée 5 (Trq-5 : +276 → +509) débutait en aval du TSS majoritaire (TSS-133117), cependant elle renfermait une boîte TAATGARAT like, un site Sp1, une boîte CAAT, une boîte TATA et deux sites c-Myb. Par ailleurs, deux troncatures de la partie 3', p-court-1 (1395pb ; -1104 → +291) et p-court-2 (687pb ; -1104 → -418), réciproques des Trq-5 et Trq-3 ont été amplifiées par PCR à partir de pGL3-p-miR-M9-M4

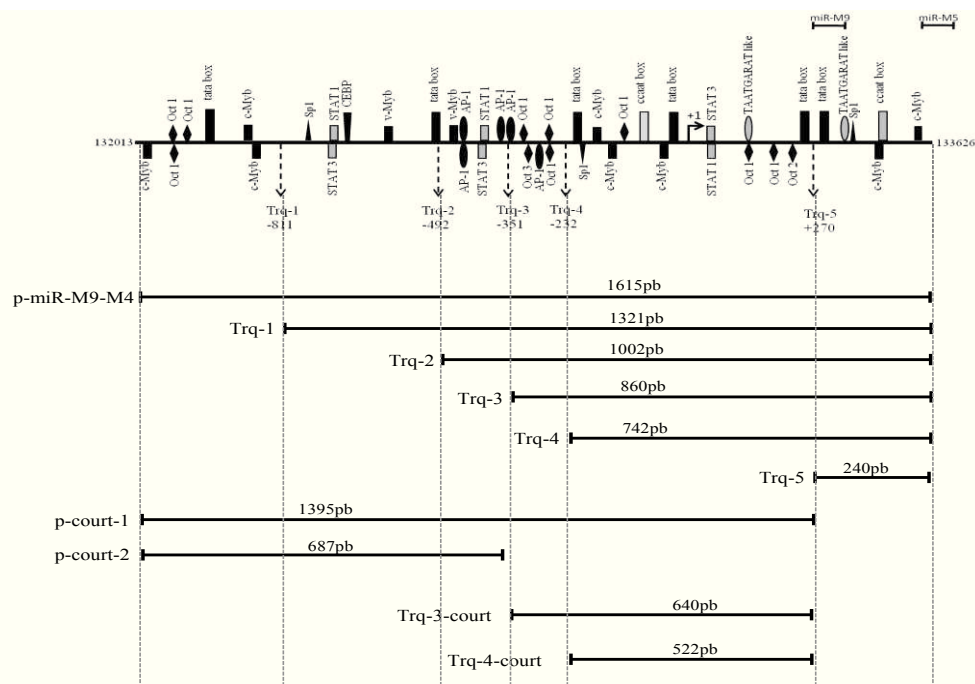


Figure 62 : Représentation schématique des différents promoteurs construits à partir de la région promotrice p-miR-M9-M4.

Les éléments de réponse aux principaux facteurs de transcription de la région promotrice p-miR-M9-M4 sont présentés en haut de la figure. La position du début des troncutures Trq-1, Trq-2, Trq-3, Trq-4 et Trq-5 est indiquée par rapport au TSS-133117 (+1). La taille de chaque promoteur est indiquée au dessus de chaque trait noir représentatif de la région correspondant aux différentes constructions.

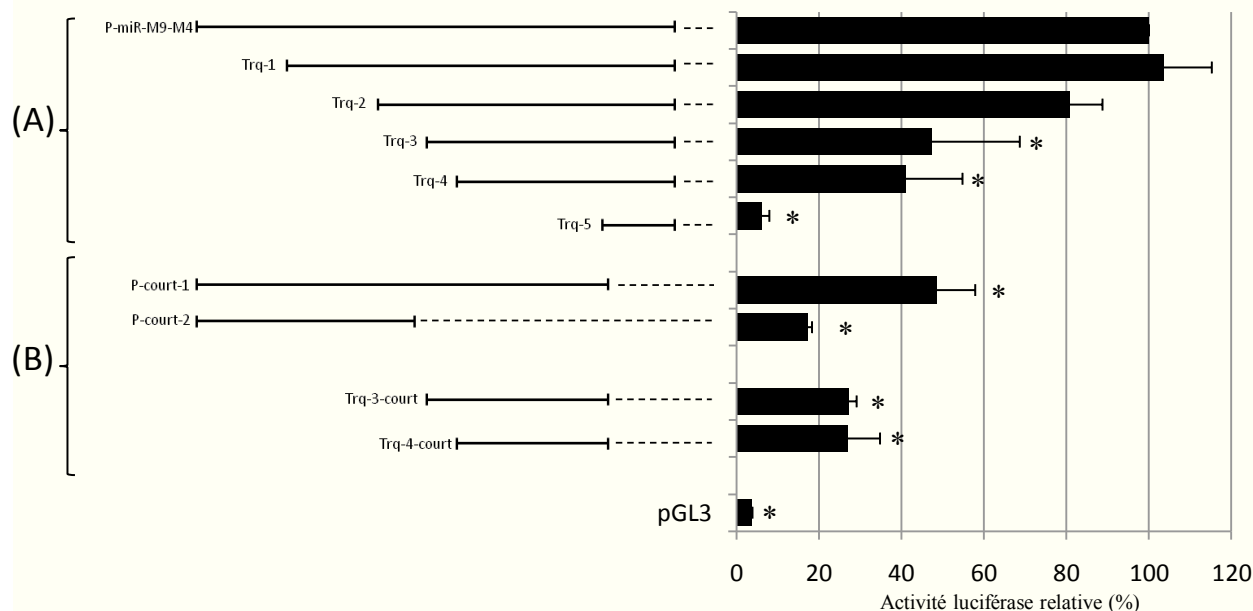


Figure 63 : Etude de l'efficacité transcriptionnelle des promoteurs tronqués de la région promotrice p-miR-M9-M4 en cellules MSB-1.

Les constructions schématisées en figure 62 sont figurées à gauche. (A) Analyse de l'activité luciférase de la région promotrice p-miR-M9-M4 et des promoteurs tronqués dans la partie 5'. (B) Analyse de l'activité luciférase des promoteurs tronqués dans la région 3' et dans les régions 5' et 3'. L'activité luciférase obtenue lors de la co-transfection de cellules MSB-1 par le vecteur pRL-TK et les différentes constructions (pGL3-Trq-1, pGL3-Trq-2, pGL3-Trq-3, pGL3-Trq-4, pGL3-Trq-5, pGL3-p-court-1 pGL3-p-court-2, pGL3-Trq-3-court et pGL3-Trq-4-court) a été relativisée par rapport à l'activité luciférase de la construction promotrice complète pGL3-p-miR-M9-M4, normalisée à 100%. Des résultats similaires ont été obtenus lors de trois essais de co-transfection indépendants réalisés en triplicate et les données présentées sont représentatives d'un essai.

respectivement à l'aide des couples d'amorces M415/M542, et M415/M476 (Figure 62 ; Tableau 22). Enfin, le panel a été complété par des formes tronquées aux 2 extrémités de la région promotrice initiale : Trq-3-court (680pb ; -351→+291), Trq-4-court (522pb ; -232→+291) et Trq-4.5-court (240pb ; +51→+291). Le mode d'obtention de ces formes a été identique à celles décrites précédemment soit des PCR à partir de la construction pGL3-p-miR-M9-M4 avec respectivement les couples d'amorces A78/ M542, M444/M542 et M477/M542 (Figure 62 ; Tableau 22).

Pour finir, toutes ces troncatures de la région promotrice p-miR-M9-M4 ont ensuite été insérées dans le vecteur pGL3 basic, afin de former les constructions pGL3-Trq-1, pGL3-Trq-2, pGL3-Trq-3, pGL3-Trq-4, pGL3-Trq-5, pGL3-p-court-1, pGL3-p-court-2, pGL3-Trq-3-court, pGL3-p-Trq-4-court, et pGL3-Trq-4.5-court.

2.3.2.2. Analyse de l'activité luciférase de différentes constructions du promoteur p-miR-M9-M4 en cellules MSB-1.

L'ensemble des constructions a été co-transfectée dans des cellules MSB-1 avec un témoin d'efficacité de transfection, le vecteur pRL-TK codant la luciférase *Renilla*. Ainsi chaque mesure d'activité luciférase Firefly a été relativisée par la valeur d'activité *Renilla* correspondante. Ensuite, afin de permettre la comparaison entre les différentes séries d'essais, les valeurs d'activité luciférase correspondant à chaque promoteur ont été standardisées en les rapportant à celles mesurées pour le vecteur pGL3-p-miR-M9-M4. Enfin, dans chaque série d'essais deux témoins ont été inclus un témoin négatif constitué par le plasmide pGL3basic sans promoteur en amont de la luciférase permettant d'évaluer le bruit de fond d'activité luciférase, ainsi qu'un témoin positif, le plasmide pCDNA-MLuc, codant pour la luciférase firefly sous promoteur CMV.

Tout d'abord, la comparaison des cinq promoteurs tronqués dans leur partie 5' a fait apparaître une activité promotrice équivalente de 100%, 105% et 80% respectivement entre les trois promoteurs p-miR-M9-M4, Trq-1 et Trq-2 (figure 63A). Par ailleurs, l'activité luciférase apparaissait similaire entre les promoteurs Trq-3 et Trq-4 et se situait à environ 45% de celle observée pour la séquence promotrice complète (Figure 63A). A l'inverse, le promoteur Trq-5 ne contenant pas le TSS majoritaire, ne présentait que 5% de l'activité de p-miR-M9-M4 (Figure 63A). Toutefois, bien que non fonctionnelle seule, la séquence de cette construction Trq-5 semble importante pour l'activité du promoteur. En effet, l'absence de cette région induit une chute d'au moins 50% de l'activité promotrice mise en évidence lors de la comparaison des constructions p-court-1 versus p-miR-M9-M4, Trq-3-court versus Trq-3 et Trq-4-court versus Trq-4 (Figure 63B). Enfin, la troncature élargie de cette région conduit à une diminution plus importante de l'activité promotrice correspondante, mis en évidence par la comparaison des activités des constructions p-court-1 et p-court-2 (Figure 63B).

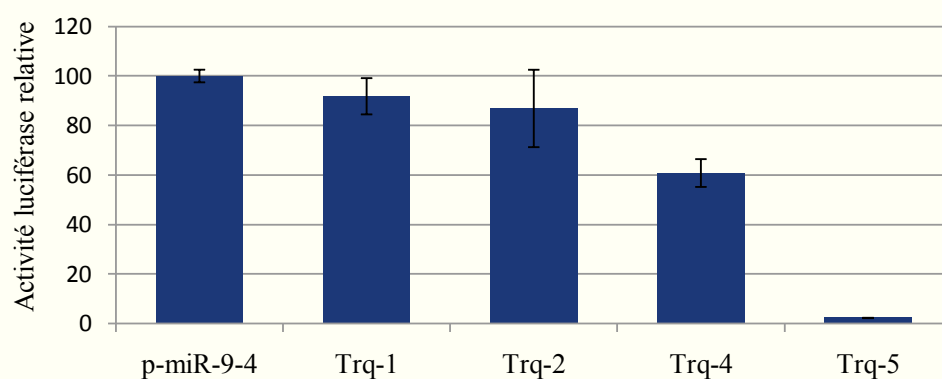


Figure 64 : Etude de l'efficacité transcriptionnelle des promoteurs tronqués de la région promotrice p-miR-M9-M4 en cellules PA-9.

L'activité luciférase obtenue lors de la co-transfection de cellules PA-9 par le vecteur pRL-TK et les différentes constructions (pGL3-Trq-1, pGL3-Trq-2, pGL3-Trq-4 et pGL3-Trq-5) a été relativisée par rapport à l'activité luciférase de la construction promotrice complète p-GL3-p-miR-M9-M4, normalisée à 100%. Les barres d'erreurs indiquent la déviation standard obtenue à partir d'un triplicate. Des résultats similaires ont été obtenus lors de deux essais de co-transfection indépendants en triplicate et les données présentées sont représentatives d'un essai.

2.3.2.3. Analyse de l'activité luciférase des différentes constructions du promoteur p-miR-M9-M4 en cellules PA-9.

Comme le promoteur présentait huit ER pour c-Myb, les constructions pGL3-p-miR-M9-M4, pGL3-Trq-1, pGL3-Trq-2, pGL3-Trq-4 et pGL3-Trq-5 ont été transfectées dans les cellules de la lignée lymphoblastoïde PA-9, qui présente un fort taux d'expression de la protéine cellulaire c-Myb suite à l'insertion du virus RAV (Rous-associated virus) dans un des allèles codant *c-Myb* (Le Rouzic & Perbal, 1996). De plus, à l'instar de la lignée MSB-1, cette lignée PA-9 est issue d'une tumeur induite par un virus GaHV-2, et dans ce cas précis la souche virulente HPRS-16.

Les résultats obtenus lors des transfections dans cette lignée surexprimant c-Myb étaient totalement superposables à ceux obtenus pour les transfections dans les cellules MSB-1 (Figure 63 et 64). En effet, l'activité luciférase relative observée pour les promoteurs Trq-1 et Trq-2 représentait environ 90% de l'activité luciférase observée pour le promoteur p-miR-M9-M4 alors que l'activité observée pour le promoteur Trq-4 ne représentait que 60% de celle observée pour p-miR-M9-M4 et que celle liée à la construction Trq-5 représentait moins de 5% de l'activité luciférase observée pour le promoteur complet.

3. Etude transcriptionnelle du brin inverse des miARN de l'IR_L de GaHV-2.

Au cours de l'étude transcriptionnelle du brin codant les miARN de l'IR_L de GaHV-2, j'ai identifié de nombreux transcrits correspondant au brin inverse (Figure 65), alors que peu de gènes avaient été caractérisés sur ce brin dans la littérature. En effet, seuls vTR localisé proche de la jonction IR_L/IR_S, la protéine 23kD et des peptides appartenant à *R-LORF10* et *R-LORF11* avaient été mis en évidence et très peu d'études avaient été menées sur les mécanismes de régulation de ces gènes.

Le tri entre les transcrits spécifiques du brin inverse aux miARN et ceux du brin sens a été réalisé sur la base de la présence de séquences spécifiques de bornes d'épissage. Ainsi, les transcrits du brin inverse localisés entre la région en amont du gène de la protéine 23kD et des répétitions de 132pb présentaient un épissage complexe de type U2 au sein desquels j'ai identifié 13 bornes accepteuses et donneuses (Tableaux 23 et 24). Par ailleurs, sur les huit ORF prédites dans cette région (*R-LORF11*, *R-LORF10*, *MRLK*, *MYFK*, *R-LORF8*, *MLHG*, *MGGG* et 23kD), les gènes *MLHG* et *MRLK* apparaissaient majoritairement exoniques tandis que les gènes 23kD, *MGGG* et *R-LORF10* semblaient principalement localisés dans des introns. Il est intéressant de noter la présence de bornes

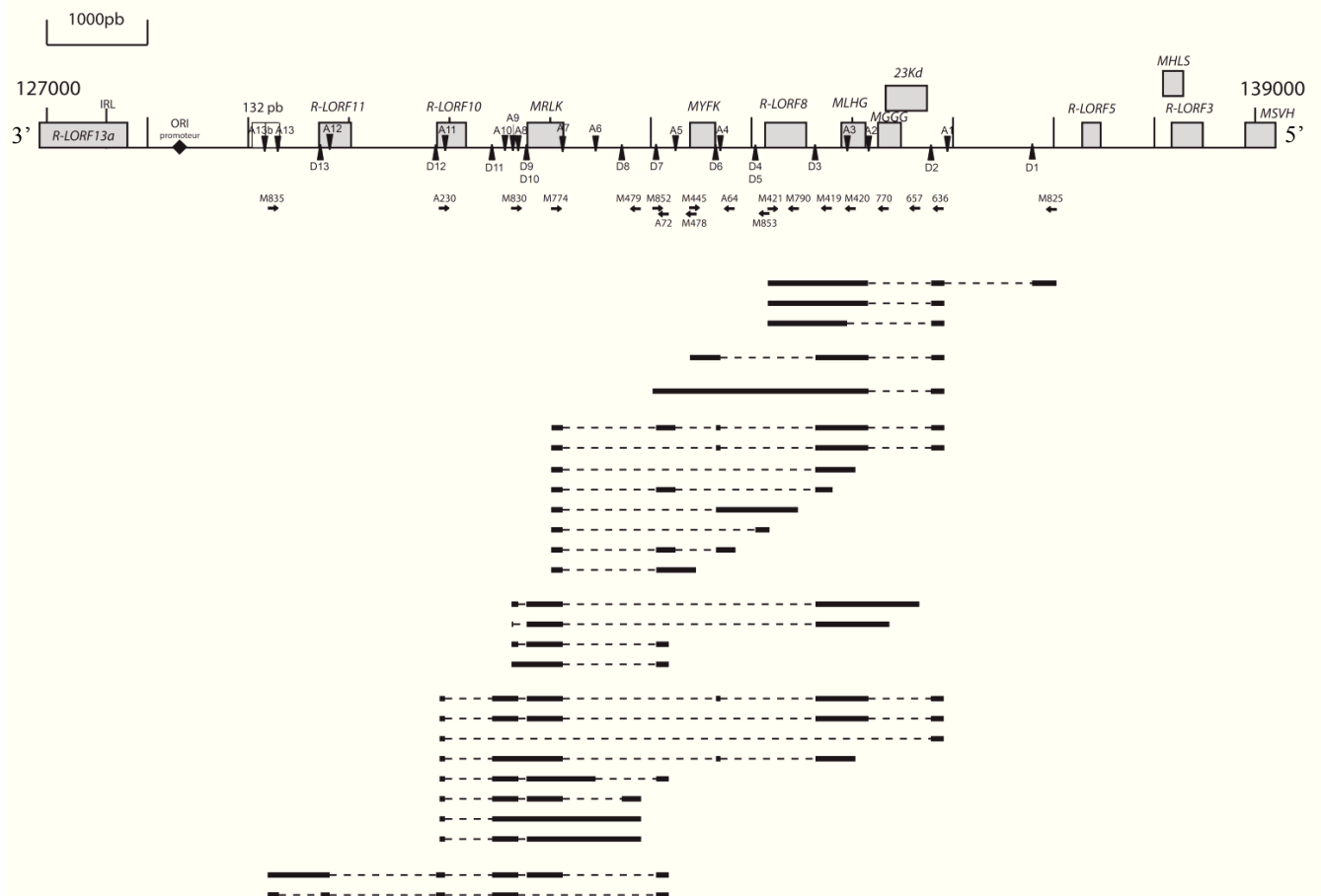


Figure 65 : Représentation schématique de l'ensemble des transcrits issus du brin inverse au brin codant les miARN du cluster1 de GaHV-2 en cellule MSB-1.

La région (3'→ 5') comprise entre les bases 127000 et 139000 de GaHV-2 Md5 est présentée dans la partie supérieure. Les ORF sont représentées par des rectangles gris. Les bornes d'épissage acceptuses de type U2 sont schématisées par des triangles noirs inversés et les bornes d'épissage donneuses de type U2 par des triangles noirs. Les amorces utilisées pour les différentes amplifications sont représentées en dessous du schéma de la région de GaHV-2. Les exons sont représentés par des traits pleins noirs, les introns de type U2 par des pointillés noirs.

	exon intron	position IR _L (Md5)
D1*	CTGCGATGAG GT ATTTTCCA [†]	136796-136777 [‡]
D2	GGGATCCTCG GT AAGACGAG	135790-135771
D3	TATTTAAGAG GT AGGTATAA	134643-134624
D4	CAGACGACAG GT ACTAAATT	134046-134027
D5	CGTGGCCACA GT AAGTCTGG	134018-133999
D6	CGATACACAC GT ATGTATTT	133655-135636
D7	TGTATGAAAG GT ATGTATTA	133060-133041
D8	GCACTATAAG GT GATTCACT	132720-132701
D9	GAAGCCTGAC GT AGGTATGA	131776-131757
D10	CCTGACGTAG GT ATGATTTT	131772-131753
D11	TTGGGGAAGG GT AAGCGAAT	131433-131414
D12	AACACGGATG GT AGGTACAC	130875-130856
D13	GCAAAAGCAG GT AAAGTATA	129729-129710

Tableau 23 : Séquence des bornes donneuses de type U2 présentes sur le brin inverse des miARN de l'IR_L de GaHV-2.

* nom des bornes D1 à D13 donné en fonction de leur position par rapport à la fin de l'IR_L (jonction IR_L/IR_S),

[†] séquence des 20 nucléotides encadrant la borne exon/intron

[‡] position des bornes donnée en fonction du génome de GaHV-2 Md5.

	intron exon	position IR _L (Md5)
A1*	ATATAAC TAG GGGAGAAGAA [†]	135952-135933 [‡]
A2	CCTTACG TAG GTGTTTCATTG	135170-135151
A3	CGTCTCG AAG TCGACCCGAG	134958-134939
A4	AACCCTG AAG ATGTTTCAAT	133699-133680
A5	ACTTCAT TAG ATCTACCCGA	133253-133234
A6	ATAATTC CAG GATCAAACCC	132459-132440
A7	CGATTTA CAG GATGAGGCTG	132135-132116
A8	GAGTTAA CAG TTCCATTTTT	131692-131673
A9	CACCGAA CAG GCGGACGGCA	131642-131623
A10	GCATGTT TAG GTCTCACCCC	131558-131539
A11	CCGTTCC TAG TTGTCGCAGT	130967-130948
A12	TCTGCCG TAG TATTTATCAT	129818-129799
A13	GTTTCGG CAG ATCGTAAGCA	129315-129296
A13b	GTTTCGG CAG ATCGTAAGCA	129183-129164

Tableau 24 : Séquence des bornes accepteuses de type U2 présentes sur le brin inverse des miARN de l'IR_L de GaHV-2.

* nom des bornes A1 à A13 donné en fonction de leur position par rapport à la fin de l'IR_L (jonction IR_L/IR_S),

[†] séquence des 20 nucléotides encadrant la borne exon/intron

[‡] position des bornes donnée en fonction du génome de GaHV-2 Md5.

accepteuses A2, A4 et A7 juste en amont des codons ATG des gènes MLHG, MYFK et MRLK, qui pourrait permettre la transcription de la totalité de ces trois ORF sous forme d'exons par épissage alternatif (Figure 65, Tableaux 23 et 24).

4. Etude de transcrits édités issus des deux brins de l'IR_L de GaHV-2.

4.1. Identification de transcrits édités au sein de l'IR_L de GaHV-2.

Au cours de l'étude transcriptionnelle de la région IR_L de GaHV-2, six transcrits fortement édités ont été identifiés à partir des ARN extraits de cellules MSB-1. En effet, de très nombreuses modifications d'adénosines en inosines (transformés en guanines lors de la transcription inverse) ont été mises en évidence sur la région de ces transcrits située juste en amont du cluster de miARN sur le brin supérieur et entre le début de R-LORF10 et l'ORF MYFK sur le brin inférieur (Figure 66).

Un seul transcrit édité (Edit-18) correspondant au brin codant les miARN a été retrouvé au sein des banques. Ce transcrit obtenu par 5'RACE PCR à l'aide de l'oligonucléotide M04, localisé dans l'exon 2 de *vIL-8* s'initiait en amont de mdv1-miR-M9 au niveau du TSS-133117 majoritaire et présentait un épissage entre la région promotrice et l'exon2 de *vIL-8*. Seul le premier exon de ce transcrit localisé en amont des miARN était édité avec 18 adénosines éditées en inosines (Figure 66). Par ailleurs, l'épissage de ce transcrit était modifié par ce phénomène d'édition qui supprimait la borne donneuse D9 et conduisait à l'utilisation d'une borne alternative située en aval.

Les cinq autres transcrits édités provenaient du brin inverse des miARN. Ils étaient tous issus de banques différentes réalisées par RT-PCR avec cinq couples d'amorces différents (Figure 66). Parmi ces cinq transcrits, différentes zones d'édition ont été mises en évidence et l'ensemble recouvrerait une zone d'environ 2500 pb du brin inverse des miARN localisée entre les gènes *R-LORF10* et *MYFK*.

Les deux transcrits Edit-175 et Edit-185 font apparaître une zone hyper-éditée entre les gènes *MRLK* et *MYFK* avec respectivement 175 et 185 modifications A→I. Les transcrits Edit-6 et Edit-88 présentait quant à eux respectivement 6 et 88 modifications A→I dans les exons présents entre et au niveau des gènes *MRLK* et *R-LORF10* (Figure 66). Le transcrit Edit-29 apparaissait non épissé et bien que recouvrant entièrement la zone entre *MRLK* et *MYFK* fortement éditée dans les transcrits Edit-175 et 185, la zone d'édition de 29 bases était située plus en amont dans la région en 3' de *MYFK* (Figure 66). Excepté pour le transcrit Edit-185 qui correspond à un épissage alternatif à partir d'une

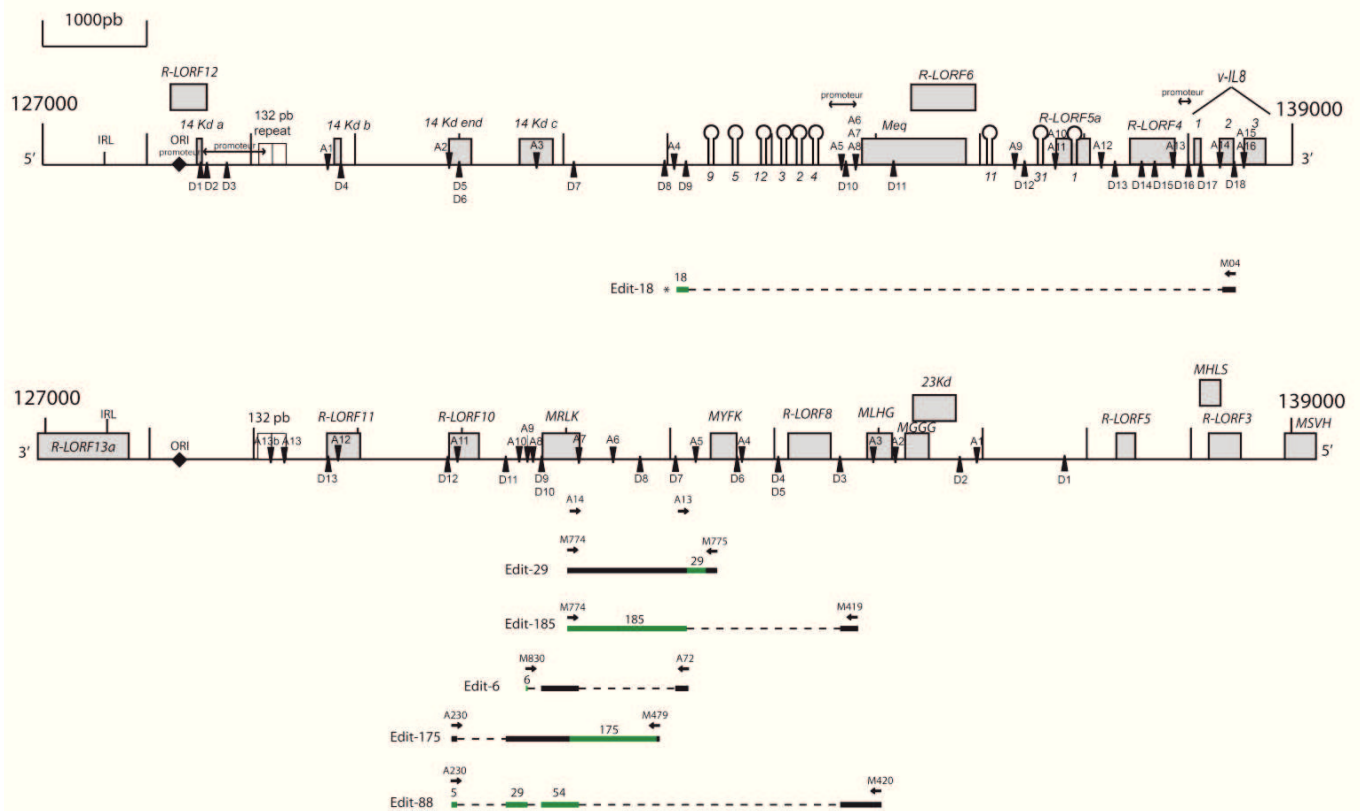


Figure 66 : Représentation schématique des transcrits édités issus de l'IR_L de GaHV-2 en cellules MSB-1.

Le brin codant les miARN de la région comprise entre les bases 127000 et 139000 de GaHV-2 Md5 est schématisé dans la partie supérieure de la figure, tandis que le brin inverse des miARN est schématisé au centre de celle-ci. Les ORF sont schématisés par des rectangles gris et les miARN par des tiges boucles. Les bornes d'épissage accepteurs de type U2 sont schématisées par des triangles noirs inversés et les bornes d'épissage donneuses de type U2 par des triangles noirs. Les exons sont représentés par des traits pleins noirs, les introns de type U2 par des pointillés noirs. Les zones éditées identifiées au sein des exons sont représentées par des traits pleins verts et le nombre de bases éditées est figuré au dessus de chaque exon édité. Les amorces utilisées pour les différentes amplifications sont représentées au dessus de chaque transcrit. L'extrémité 5' identifiée par 5'RACE PCR est indiquée par une étoile (*).

borne accepteuse créée par édition (intron/exon AA/AT en AG/AT), les bornes d'épissage des autres transcrits édités sont identiques à celles précédemment décrites pour les transcrits non édités (Figure 65). En plus de la génération d'un site d'épissage alternatif, les modifications les plus importantes lors de l'édition du transcrit Edit-185 concernaient les codons stop (TAA, TAG et TGA) avec la disparition de 36 des 62 initialement présents sur l'exon 2 de ce transcrit (Figure 67). De plus, l'édition induisait la suppression de quatre codons d'initiation ATG et la création de 8 ATG à l'intérieur de ce second exon.

Il faut noter que la création de ces codons ATG et le « gommage » des nombreux codons d'arrêt de traduction au sein de l'exon 2 du transcrit Edit-185 ont entraîné une importante modification des cadres ouverts de lecture de cette région (Figure 68). En effet, l'édition du transcrit Edit-185 a fait apparaître trois petits ORF de 84, 96 et 123 nucléotides, ainsi que l'allongement de deux cadres de lecture existants. Le premier ORF concerné situé dans la première phase de lecture comporte de ce fait 36 nucléotides supplémentaires alors que l'apparition d'un ATG dans la deuxième phase de lecture permet l'adjonction de 519 bases (soit 173 aa) en 5' du cadre ouvert de lecture MRLK transformé ainsi en MRLE (Figure 68).

4.2. Etude de la variabilité de l'édition dans la région Edit-29.

En préambule à cette dernière partie des résultats, je voudrais indiquer qu'étant à l'origine de la mise en évidence des transcrits édités au sein de l'équipe, nous avons tenu à insérer dans ce manuscrit, la partie des analyses de ces transcrits que j'ai réalisée. Ces résultats obtenus en fin de stage de thèse, ayant été uniquement initiés en vue d'évaluer la faisabilité des approches, sont naturellement parcellaires et préliminaires. Par contre, ils sont à l'origine du projet de recherche d'un étudiant en M2, qui a pour but de compléter les profils des transcrits édités et surtout de déterminer l'enzyme impliquée dans ce mécanisme d'édition.

Afin de déterminer l'existence potentielle de plusieurs types de profils d'édition et surtout si certains événements d'édition étaient particulièrement présents au sein d'une même zone, nous avons réalisé des RT-PCR et RT-PCR nichées à partir d'ARN extraits de cellules MSB-1 en ciblant la région éditée du transcrit Edit-29. Tout d'abord, l'étape de transcription inverse a été réalisée à l'aide d'un mélange de deux oligonucléotides A13 et A14, spécifiques du brin inverse des miARN. Ces amorces localisées respectivement 90pb en aval de la région éditée et dans l'ORF de MRLK souvent retrouvée dans un exon, permettaient d'obtenir les ADNc des différents produits d'épissage identifiés parmi les transcrits édités (Figure 66). Afin d'amplifier les événements d'édition, des PCR ont ensuite été réalisées à l'aide de quatre couples d'amorces (M1 à M4) correspondant à l'association des amorces spécifiques des séquences éditées A5-A7 (M1), A5-A6 (M2), A8-A7 (M3) et A8-A6 (M4) alors que les PCR témoins équivalentes étaient réalisées avec les couples d'amorces « non éditées » A9-A11 (M5), A9-A10 (M6), A12-A11 (M7) et A12-A10 (M8) (Figure 69A et B). Ensuite tous les produits

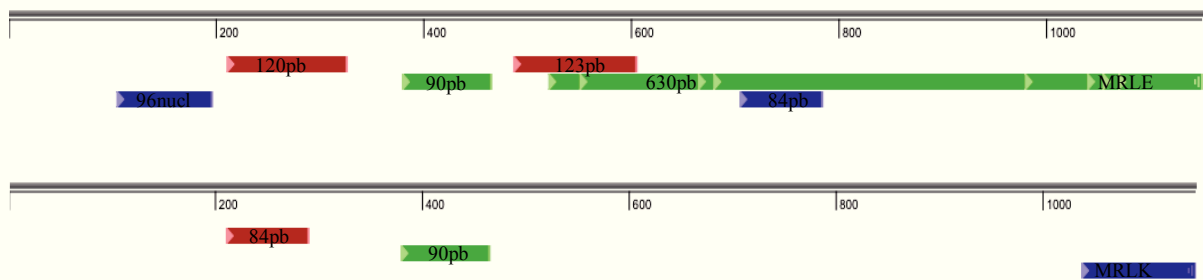
Edit-185	1	C	TAGGAATCAGTGTGCGGAATTTATCT	TAA	CATTCCAGCACCAACCTCCCCGAACCAAA	60							
Md5	134806	C	TAGGAATCAGTGTGCGGAATTTATCT	TAA	CATTCCAGCACCAACCTCCCCGAACCAAA	134747							
Edit-185	61	A	TAATAA	TAA	AAAAGCAACACCCACAGACCCGAAAGTAT	TGA	AGAATTACACGTTTCGCG	120					
Md5	134746	A	TAATAA	TAA	AAAAGCAACACCCACAGACCCGAAAGTAT	TGA	AGAATTACACGTTTCGCG	134687					
Edit-185	121	A	ACGGTCACAATTCACCTGTCATTTTA	TAA	AGGAGCAATAG	TTTATT	TAA	GAG	173				
Md5	134686	A	ACGGTCACAATTCACCTGTCATTTTA	TAA	AGGAGCAATAG	TTTATT	TAA	GAG	134634				
Edit-185	174	ATG	TCGCGGTCGAGGGGCCGGCTGGTGGAGGGATCGGGCGGGGAATCGGGGAGCGGACG	233									
Md5	133161	ATG	TCACGATCGAAAAGCCGGCTGG	TAA	AGGGATCGGGCGAAAAGATCGGGGAGCGGACG	133102							
Edit-185	234	GACT	TGA	GCCTTGGGTGGAG	TGA	CTGGCGCGCAGCCAGATCCTGT	AT	GAAAGGTGTGTGTT	293				
Md5	133101	GACT	TGA	GCCT	TAA	G	TAA	CTGGCGCGCAGCCAGATCCTGT	AT	GAAAGGT	ATG	TATT	133042
Edit-185	294	G	GGAGTTCGATTGGCGGAACAGGGCAATTCGCGCAGGATT	ATG	CGTAA	TCAAGCCACAG	353						
Md5	133041	A	GAAAATTCGATTGGCGGAACAGGGCAATTCGCGCAGGATT	ATG	CGTAA	TCAAGCCACAG	132982						
Edit-185	354	A	TCGCACCTTCAACCGA	TAA	GAGTCTC	TAG	AT	GACAAATCGACAGT	TGGTTGGAGGGTGT	413			
Md5	132981	A	TCGCACCTTCAACCGA	TAA	GAGTCTC	TAG	AT	GACAAATCGACAGT	TAA	TGGAAAGTGT	132922		
Edit-185	414	T	CGGGGGCGG	TAA	AATTTGCGGGGCTCTCGACATAA	AGGCCCTCTGGGCT	ATG	GTTTCGCT	473				
Md5	132921	T	CGGGGGCGG	TAA	AATTTACAGGGCCTCTCGACATAA	AGGCCCTC	TAA	AGTA	TAA	TTCGCT	132862		
Edit-185	474	G	GTTCCTG	TGA	TTGGGGGAGCGGTTCT	TGA	TGGTTCGTACACCCGATCGCATATTT	ATG	C	533			
Md5	132861	G	GTTCCTG	TAA	TAG	GAAGAGCGGTTCT	TGA	C	TTCGTACACCCGATCGCATATTT	ATG	C	132802	
Edit-185	534	T	GGGGCGAGGGGTGG	TGA	G	ATG	GTGGGTTCGGCACT	TAA	GAAAGGGCGTCTGTCTGTACAG	593			
Md5	132801	TGA	GGCGAGGGG	TAA	TGA	A	TGATA	ATCAGCACT	TAA	GAAAGGGCATCTGTCTATCAG	132742		
Edit-185	594	G	CCCCGGACATCAAAGGCACTGCACATA	TAG	GGTGA	TTCACTTGGGC	TGA	GATTCCGGGCA	653				
Md5	132741	G	CCCCGAGACATCAAAGGCACTGCACATA	TAA	GGTGA	TTCACT	TAA	GCTAA	AATTCGGGCA	132682			
Edit-185	654	T	CGGCT	ATG	SGGTGGGCAATATCAGTTT	TCG	GGGGGTTTCTC	ATG	GGAACAGCGACTTGGGA	712			
Md5	132681	T	CAGCT	AT	T	G	TAA	GCAATATCAGTTATCAAAATTTCTC	ATG	GGAACAGCGACT	TAA	AA	132622
Edit-185	713	A	-GAACCTTGA	TAT	TGGAGGGGTGCCTTCAGTGCCCG	ATG	CGGTTGTGGACGGTTC	CCGTG	772				
Md5	132621	A	AAACCTTAT	TAA	TAA	AAGGAAATACCTTCAATACCCA	ATG	CGTTA	TAA	ACAATTC	CCGTA	132562	
Edit-185	773	T	AGGTCC	TGA	GGAGCTGCCGACGGAAGAGGAAAT	GCCCCATTT	CGGCCATTGGAGCGT	832					
Md5	132561	T	AAATCC	TGA	GGAGCTGCCGCAAAAAA	---CTACCCCATTT	CGGCCATTGGAACT	132505					
Edit-185	833	G	TGGA	ATG	GCAACAACAAAA	ATG	GGGGCGCAGTCGGCACGGTTTC	ATG	GTTCCGGGATCG	892			
Md5	132504	A	TAA	AA	TAA	CAACAACAAAA	TAA	AAACGCAGTCGGCACGGTTTCA	TAA	TCCAGGATCA	132445		
Edit-185	893	G	GGCCGGCCGGGCGCTCTCGAGGTGTGGCACACCCCTCCCCGCGGCTGCTGGTCCGGA	952									
Md5	132444	A	ACCCGGCCGAACGCTCTCGAATA	TAA	CACACCCCTCCCCGCGGCTACTGGTCCGGA	132385							
Edit-185	953	G	GTCAT	TGA	GTTTGGTGGCAGACGACAAACAGACAAGGTTGTGG	ATG	CCGCCGGCTGGT	1012					
Md5	132384	A	ATCAT	TGA	GTT	TAG	TGGCAACGACAAACAGACAAAAATG	TAA	ATG	CCGCCAGC	TAA	T	132325
Edit-185	1013	T	GAGGGGCGCACGGAACGTGTGTGCGGGCGCGTGGAGCAACCGACACCGAGTTGGTTT	1072									
Md5	132324	TGA	GAGCGCACAGAATATATCGAGACG	TAA	TAGCAACCGACACCGCAAGT	TAA	TTT	132265					
Edit-185	1073	A	CGGCGGTTGCTGTGCGTTTGCAG	TGA	CTTTTTTT	AG	CCT	TGATCGTGCCGTGCCCGGG	1132				
Md5	132264	A	CGGCGGTTACTGTACATTTCAG	TAA	CTTTTTT	TAA	CCT	TGATCGTACCGTACGCGAAG	132205				
Edit-185	1133	G	GGAGTGTGCGAGGGGTGGA	ATGA	GCGCGCTACAGTGGGCCAGTGTCCGTGGCCTTTTG	1192							
Md5	132204	G	GGAGTATACGAAAGA	TAA	AATGA	ACACGCATACAG	TGA	ACCAATATCCG	TAA	CCTTTTA	132145		
Edit-185	1193	G	TGCTCGACCGATTTCGAGG	ATGA	GGCTGGAAGT	ATG	CATTCGGA	ATG	GATTGGAA	1252			
Md5	132144	ATG	CTCGACCGATTTCAGAG	ATGA	GGCTGGAAGT	TGA	AAAATATACATACATTCAAA	TAA	TAA	AA	132085		
Edit-185	1253	A	AGGCGGAGACG	ATG	GGCGTCCCCGCTGCACGTCCGCGTTTCGTCTCGGCGCCGCGTACT	1312							
Md5	132084	A	AAACAAAGACA	TAA	ACGTCCCCGCTACACGTCCGCGTTTCGTCTCGGCGCCGCGTACT	132025							
Edit-185	1313	T	CCGCAATCGTT	1324									
Md5	132024	T	CCGCAATCGTT	132013									

Exon 1
non édité

Exon 2
édité

Figure 67 : Alignement de la séquence Edit-185 avec la région correspondante du génome GaHV-2 Md5. L'alignement entre la séquence Edit-185 et le génome de GaHV-2 Md5 a été réalisé à l'aide du logiciel BLAST. Les codons ATG non modifiés par le phénomène d'édition sont surlignés en jaune, ceux créés par édition sont surlignés en vert clair et ceux supprimés par l'édition sont surlignés en rouge. Les codons stop (TAA, TAG et TGA) non modifiés par le phénomène d'édition sont surlignés en gris et ceux supprimés par l'édition sont surlignés en vert foncé.

Séquence éditée Edit-185



Séquence non éditée

Figure 68 : Prédiction des cadres ouverts de lecture de l'exon édité de la séquence Edit-185 en comparaison de la séquence non éditée de GaHV-2 Md5.
Les trois cadres ouverts de lecture sont représentés en rouge, vert et bleu.

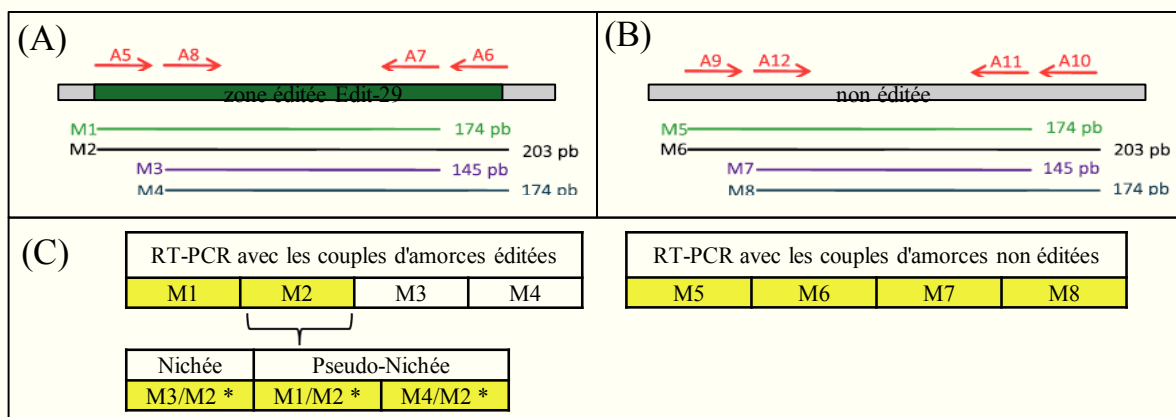


Figure 69 : Descriptif des différentes RT-PCR et RT-PCR nichées réalisées dans la région éditée du transcrit Edit-29 en cellules MSB-1.

(A) Couples d'amorces édités utilisés pour les RT-PCR et RT-PCR nichées ainsi que la taille des fragments attendue. (B) Couples d'amorces non édités utilisés pour les RT-PCR et RT-PCR nichées ainsi que la taille des fragments attendue. (C) Ensemble des RT-PCR et RT-PCR nichées réalisées à partir des ADNc issus des transcriptions inverses spécifiques du brin anti-sens des miARN. Les cases surlignées en jaune signifient qu'une amplification a été observée sur gel. Les étoiles correspondent aux produits de PCR au sein desquels des transcrits édités ont été identifiés. La notation Mx/My symbolise une RT-PCR nichée ou pseudo-nichée avec le couple d'amorce x sur le produit de RT-PCR ayant utilisé le couple d'amorces y.

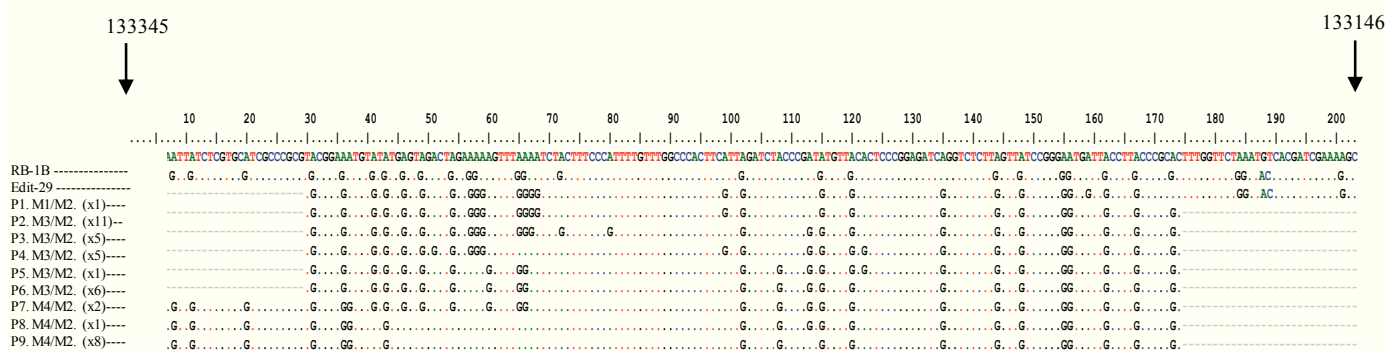


Figure 70 : Alignement des différentes séquences éditées.

Les séquences éditées sont alignées grâce au logiciel BIOEDIT en fonction du génome viral Md5 et de la séquence éditée Edit-29. Les neuf populations retrouvées (P1 à P9) ainsi que les pcr dont elles sont issues sont indiquées sur la droite de la figure. Le nombre d'éléments de chaque population est indiqué entre parenthèses. Les points montrent les bases identiques par rapport à la séquence Md5 et les tirets des « gaps ».

d'amplification ont été clonés en vecteur pGEM-T Easy afin de déterminer les séquences des amplicons correspondants. Alors que les PCR témoins avec les amorces non éditées ont permis l'obtention d'amplicons spécifiques, la totalité des amplicons obtenus pour les couples M1 à M4 correspondait à l'amplification et au clonage de séquences d'ARNm de poulet dont les séquences étaient très proches des séquences des amorces « éditées ». Pour augmenter la spécificité des amplifications lors de ces PCR, des PCR nichées ont été réalisées avec le couple M1, M3 ou M4 sur le produit de PCR amorcé par M2 (Figure 69C).

Des amplicons de taille attendue ont été obtenus pour chaque PCR nichée. Après purification sur gel, les ADN correspondants ont été clonés en vecteur pGEMTeasy. Suite au séquençage, 28 séquences éditées de la région correspondant à la PCR nichée M3 à partir de la PCR primaire M2 (M3/M2) ont été obtenues alors qu'une et onze séquences correspondant respectivement aux PCR nichées M1/M2 et M4/M2 ont été trouvées.

L'alignement des 40 séquences réalisé à l'aide du logiciel BIOEDIT a permis de regrouper les séquences identiques entre elles (Figure 70) et de répertorier neuf profils d'édition (P1 à P9). Ainsi, le profil majoritaire correspondait au profil P2 (11 clones), les profils P9, P6, P3 et P7 concernant respectivement 8, 6, 5 et 2 clones alors que les profils P1 et P8 n'avaient été rencontrés qu'une seule fois sur les 40 séquences déterminées (Figure 70).

L'analyse globale de l'alignement montre que 42 adénosines sur les 55 que contenait la région amplifiée pouvaient être éditées en inosines. Par ailleurs, il est apparu que plusieurs d'entre elles, localisées en position 31, 36, 43, 102, 120, 144, 148, 155, 156, 162 et 167, étaient systématiquement éditées. De plus, il semblerait par la comparaison des populations que certains sites d'édition soient favorisés ou bloqués par l'édition d'un autre site. En effet, l'édition du site en position 108 pour les populations P5 à P9 et l'édition du site en position 60 identifiée dans les populations P5 à P7 semblaient exclusives de l'édition sur les sites 57 à 58 observée sur les populations P1 à P4 ainsi que sur le transcrit Edit-29 (Figure 70). Par ailleurs, pour la population P4, l'édition en position 51 semblait exclure l'édition en positions 65 et 66 présente dans les populations P1 à P3. Pour finir, l'édition observée pour les bases 57-59 et 65-68 semblerait montrer que l'édition d'une adénosine au sein d'une série d'adénosines favoriserait l'édition des bases situées immédiatement en aval (Figure 70).

DISCUSSION

Le miARN mdv1-miR-M4 localisé dans le cluster M9-M4 en amont du gène *meq* apparaît comme un des miARN les plus exprimés durant l'infection par GaHV-2. En effet, parmi les différentes banques de miARN réalisées à partir de lignées de tumeurs induites par GaHV-2, d'infection *in vitro* de CEF par GaHV-2 ou *in vivo* de PBL de poulets infectés par GaHV-2, mdv1-miR-M4-5P correspondait à l'un des trois miARN les plus exprimés, représentant de 6 à 72% des miARN viraux clonés dans les différentes banques. Parallèlement à nos travaux, ce miARN avait été identifié comme un orthologue du miARN cellulaire gga-miR-155 sur la base de l'identité de leur séquence « seed » (UAAUGCU). Ainsi, il avait été montré que ces miARN diminuaient identiquement l'expression des gènes rapporteurs contenant les 3'UTR des gènes *PU.1*, *CEBPβ*, *HIVEP2*, *BCL2L13* et *PDCD6*, ainsi que l'expression de la protéine PU.1 (Zhao *et al.*, 2009a).

Dans ce travail, nous avons montré un effet de gga-miR-155 et de son orthologue mdv1-miR-M4-5P sur trois nouvelles cibles cellulaires RREB1, c-Myb et GPM6B ainsi que sur trois cibles précédemment identifiées pour hsa-miR-155 (*MAP3K7IP2*, *PU.1*, *CEBPβ*) (Ceppi *et al.*, 2009; Vigorito *et al.*, 2007; Yin *et al.*, 2008) parmi lesquelles le ciblage de deux d'entre elles avait parallèlement été décrit pour gga-miR-155 et mdv1-miR-M4-5P (*PU.1*, *CEBPβ*) (Zhao *et al.*, 2009a). Alors qu'une répression identique a été observée pour les vecteurs véhiculant les 3'UTR des gènes *GPM6B*, *CEBPβ*, *MAP3K7IP2*, *RREB1* et *PU-1* avec les deux miARN (Figure 48), une légère différence a été observée pour la cible c-Myb présentant deux miRE. En effet, mdv1-miR-M4 semble diminuer plus efficacement l'expression de c-Myb que gga-miR-155 via le miRE1 (Figure 49B) avec de plus, un potentiel effet coopératif lorsque les deux miRE sont présents lors de l'expression ectopique de mdv1-miR-M4. Il faut noter que la différence d'activité luciférase observée entre mdv1-miR-M4-5P et gga-miR-155 sur le miRE1 de c-Myb (mutation miRE2 ; Figure 49B) ne semble pas pouvoir être reliée à une meilleure complémentarité entre le mdv1-miR-M4-5P et la cible (Figure 47). Par contre, cet effet coopératif entre les deux miRE des gènes c-Myb et RREB1 est nettement détectable avec mdv1-miR-M4 exprimé de façon endogène lors des essais dans les lignées lymphoïdes MSB-1 et 54-O. En effet, la présence des deux miRE au sein de la 3' UTR du messager de RREB1 entraîne une diminution de l'activité luciférase d'environ 90%, alors que séparément les deux miRE1 et 2 ne diminuaient l'activité que de respectivement 50% et 70 % (Figure 50). De la même manière, une diminution cumulée de l'activité luciférase d'environ 70% a été observée pour c-Myb alors que séparément chacun des deux miRE n'entraînait qu'une diminution de l'activité luciférase de 50% (Figure 50).

Par ailleurs, l'alignement dans différentes espèces des sites miRE2 de c-Myb et de RREB1 présentant une grande variabilité sur l'ensemble de la séquence du miRE ainsi que des mutations dans la séquence complémentaire à la séquence « seed » chez le taureau (bta) et le chien (cfa), indiquerait que ce second site ne serait pas fonctionnel chez toutes les espèces (Tableau 25). A l'inverse, la remarquable conservation des séquences et de leur localisation au sein des 3'UTR chez les différents organismes laisse supposer que les sites miRE1 de RREB1 et de c-Myb seraient systématiquement

cible	espèce	taille 3'UTR	miRE	distance miRE/polyA	séquence (5'→3')*
RREB1	gga [#]	1710	1	1461	AGGACC <u>UUUUUCAUUUA</u> AGCAUUAA
	hsa	1994	1	1652	GAGGACC . <u>UCUUCAUUU</u> GAGCAUUAG
	mml	1923	1	1586	GAGGACC . <u>UCUUCAUUU</u> GAGCAUUAG
	mmu	1864	1	1546	GAGGACC . <u>UUUUUCAUUUA</u> AGCAUUAA
	rno	1823	1	1512	GAGGACC . <u>UUUUUCAUUUA</u> AGCAUUAA
	bta	1828	1	1533	GAGGACC . <u>UUUUUCAUUUA</u> AGCAUUAA
	cfa	1851	1	1486	GAGGACC . <u>UUUUUCAUUUA</u> AGCAUUAA
	gga [#]	1710	2	1226	<u>CAUUUGCCACACGGUGG</u> AGCAUUAA
	hsa	1994	2	1306	<u>UCCAUUU . CCAGUCUGU .</u> CAGCAUUAG
	mml	1923	2	1240	<u>UCCGUUU . CCACUCUGU .</u> CAGCAUUAG
c-Myb	bta	1828	2	1174	<u>CUCUCAGCGUAGACAUG</u> CCAUAUAC
	gga [#]	1211	1	1178	AUGGAC <u>UUCUCCGGAGA</u> AGCAUUAA
	hsa	1191	1	1169	GACAUUUCCAGAAA AGCAUUAA
	mml	1188	1	1166	GACAUUUCCAGAAA AGCAUUAA
	mmu	1232	1	1210	GACAUUUCCAGAAA AGCAUUAA
	rno	1254	1	1232	GACAUUUCCAGAAA AGCAUUAA
	bta	1199	1	1177	GAC <u>GUUU</u> GAGCAACA AGCAUUAA
	cfa	1174	1	1152	GA <u>UGUUUC</u> UACAACA AGCAUUAA
	gga [#]	1211	2	734	<u>AAAUACUGAAGUAAAA</u> AGCAUUAA
	hsa	1191	2	753	UUUUAAAAAUCAGUAAA AGCAUUAA
PU.1	mml	1188	2	754	UUUUAAAAAUCAGUAAA AGCAUUAA
	mmu	1232	2	782	UUUUAAAAA CCAGGAGA AGCAUUAA
	rno	1254	2	810	UUUUAAAAA CCAGUAAA AGCAUUAA
	bta	1199	2	777	<u>UCAU</u> UACAUAAA . CAC . . AA UGCAUUAA
	cfa	1174	2	761	<u>UCA</u> ACCACUAAA . CAC . . AA UGCAUUAA
	gga [#]	427	2	94	<u>CUUACAGCUUUCAAACU</u> AGCAUUAA
	hsa [#]	368	1	316	<u>GCCUCCCCGUGGCCAU</u> AGCAUUAA
	mmu [#]	357	1	308	<u>UGGACCCCGCCGCCAU</u> AGCAUUAA
	rno	364	1	315	<u>CGGCCCCCGCCGCCAU</u> AGCAUUAA
	bta	422	1	368	<u>GCCGCCCCGUGGCCAU</u> AGCAUUAA
MAP3K7IP2	cfa	405	1	347	<u>GCCUCCCCCGCCGCCAU</u> AGCAUUAA
	gga [#]	1857	1	1031	CAAUGCCUU . AAAAU . . . AA GAGCAUUAA
	hsa [#]	1829	1	1004	AUGCCUU . AAAAU . AAAA AGCAUUAA
	mml	1908	1	1081	AUGCCUU . AAAAU . AAAA AGCAUUAA
	mmu	1780	1	1036	AAUGCCAU . AAAGGU . . AGAA AGCAUUAA
	rno	1799	1	1041	UGCCAU . ACGAUAGAAA GAGCAUUAA
	bta	1888	1	1062	UGCCU U CAAAAU . AAAA AGCAUUAA
	cfa	1917	1	1088	<u>UGCCUU . AAAAU . AAAA</u> AGCAUUAA
	CEBPβ	315	1	21	<u>AUUCUUGCUACA</u> UUUGA AGCAUUAA
	hsa [#]	604	1	43	GAAU CUUUUCCGUU UCA AGCAUUAA
GPM6B	mml	606	1	47	<u>GGGAU</u> CUUCCGUU UCA AGCAUUAA
	mmu [#]	505	1	45	GAAC CUUUUCCGUU UC GAGCAUUAA
	rno	487	1	46	GAAC CUUUUCCGUU UC GAGCAUUAA
	cfa	642	1	58	<u>GAU</u> CUUUUCCGUU UCA AGCAUUAA
	gga [#]	1713	1	155	AAUA . <u>UUUUUAUGUAAAG</u> AGCAUUAA
	hsa	1937	1	203	AUAUUUUUAUGUAAAG AGCAUUAA
	mml	1886	1	223	AUAUUUUUAUGUAAAG AGCAUUAA
	mmu	1879	1	204	AUAUUUUUAUGUAAAG AGCAUUAA
	rno	1868	1	204	AUAUUUUUAUGUAAAG AGCAUUAA
	bta	1902	1	204	AUAUUUUUAUGUAAAG AGCAUUAA
	cfa	1939	1	203	AUAUUUUUAUGUAAAG AGCAUUAA

Tableau 25 : Alignement des miRE ciblés par miR-155 et son orthologue mdv1-miR-M4-5P.

[#] Cibles déjà validées pour le miR-155 de l'espèce correspondante, gga : *Gallus gallus* ; hsa : *Homo sapiens* ; mml : *Macaca mulatta* ; mmu : *Mus musculus* ; rno : *Rattus norvegicus* ; bta : *Bos taurus* ; cfa : *Canis familiaris*

* Les nucléotides soulignés en gras correspondent à la séquence complémentaire de la « seed » de miR-155, les bases en rouge correspondent à des mutations et les points représentent des « gaps ».

impliqués dans la régulation de l'expression des deux gènes correspondants au cours de l'évolution (Tableau 25).

Pour la cible PU.1, en plus du miRE, miRE2 déjà identifié dans la littérature (Zhao *et al.*, 2009b), un second miRE miRE1 avait été prédit par recherche bioinformatique spécifiquement pour le gène du poulet. De plus, la courte distance séparant ces deux miRE semblait en faveur d'un potentiel effet synergique (Grimson *et al.*, 2007). Cependant, ce site potentiel miRE1 qui ne présentait pas un appariement canonique avec les nucléotides 2 à 7 de la séquence « seed » s'est avéré non fonctionnel pour les deux miARN orthologues (Figure 49), définissant ainsi le miRE2 comme l'unique élément de réponse à ces deux miARN. De plus, il faut noter que contrairement aux autres miRE testés au cours de ma thèse (notamment pour RREB1, c-Myb, MAP3K7IP2, CEPB β et GPM6B), la localisation de ce miRE à 94 nt de l'extrémité 3' du messager est différente chez le poulet de celle observée (environ 300 nt) chez toutes les autres espèces répertoriées (Tableau 25). Cette observation est à rapprocher de ce qui a été décrit pour le gga-miR-150 pour lequel la régulation du gène c-Myb de poulet s'effectuait au travers du miRE orthologue du poisson zèbre (*Danio rerio*), plutôt qu'au travers des deux miRE communs avec les mammifères (Guillon-Munos *et al.*, 2010). Enfin, l'analyse fine des séquences des différents miRE de cette étude montre que, hormis pour le miRE de GPM6B révélant une très grande conservation de l'ensemble de la séquence du miRE, la partie 5' des miRE correspondant à la région 3' du miR-155, identifiées chez le poulet diffèrent de celles des mammifères.

Globalement, nos résultats ont contribué à accroître le nombre de cibles identifiées pour miR-155 et son orthologue mdv1-miR-M4-5P confortant, entre autres, l'impact potentiel de ces miARN dans la différenciation des lymphocytes et la réponse immunitaire via un effet indirect sur l'expression de différentes interleukines.

Plus précisément, les protéines PU.1 et c-Myb ont été caractérisées comme des régulateurs principaux du développement et de la différenciation des lymphocytes B et T. Le facteur de transcription PU.1 est essentiel pour la différenciation terminale des cellules myéloïdes, le développement des lymphocytes B et T, l'érythropoïèse et la maintenance des lignées hématopoïétiques (Kastner & Chan, 2008). Le facteur de transcription c-Myb a été montré essentiel lors de l'établissement du système hématopoïétique et durant plusieurs étapes du développement des lymphocytes T. Par ailleurs, c-Myb a été identifié récemment comme un transactivateur de hsa-miR-155 par fixation au sein de la région promotrice de ce dernier, créant potentiellement une boucle de régulation négative entre le miARN et sa cible (Vargova *et al.*, 2011). Un des rôles primordiaux de miR-155 et de ses orthologues semble correspondre à la régulation de la réponse immunitaire par régulation indirecte de l'expression des interleukines. En effet, plusieurs cibles directes de ces miARN sont impliquées dans la régulation d'IL-1, IL-4, IL-5, IL-6, IL-8, IL-10, IL-13, MIP-1 α (macrophage-inflammatory protein-1alpha) et du TNF α . Par ailleurs, dans le contexte d'une infection par GaHV-2, il est intéressant de noter que la balance entre la latence et la réplication est en partie contrôlée par la production de cytokines (Schat & Nair, 2008a). Ainsi, mdv1-miR-M4-5P pourrait permettre le

contrôle du cycle viral de GaHV-2 par action indirecte sur les cytokines. Ce rôle concorderait avec les taux de cytokines observés lors de la pathogenèse de GaHV-2 chez le poulet. En effet, il a été décrit au début de l'infection par GaHV-2, lors de la première phase cytotytique, une augmentation du taux de nombreuses interleukines puis une diminution du taux de ces dernières lors de la phase de latence/oncogenèse (Abdul-Careem *et al.*, 2008; Heidari *et al.*, 2008). Plus précisément, les facteurs IL-1 β , IL-4, IL-6, IL8, TNF α et MIP-1 α sont directement régulés par CEBP β , cible de miR-155 testée au cours de notre étude (Akira *et al.*, 1990; Davydov *et al.*, 1995; Kunsch *et al.*, 1994; Pope *et al.*, 1994; Rosati *et al.*, 2001; Zhang & Rom, 1993), tandis que la protéine MAP3K7IP2 agirait sur la cascade de signalisation TLR/IL-1 et son inhibition par miR-155 entraînerait une diminution de la production d'IL-1 β et de TNF α (Ceppi *et al.*, 2009). De plus, la protéine IL-1 est également impliquée dans l'activation de NF κ B (nuclear factor kappaB) et de MAPK8 (JNK). Il faut noter que MAPK8 qui joue un rôle clef dans la prolifération des lymphocytes T, l'apoptose et la différenciation est également activée par le facteur MAP3K10, lui même régulé par miR-155 (Gottwein *et al.*, 2007). Par ailleurs, Rodriguez et ses collaborateurs ont montré une diminution de l'expression de c-Maf par miR-155 ; or l'expression ectopique de c-Maf est capable d'induire l'expression des cytokines IL-4, IL-5 et IL-10 dans des cellules Th2 (Rodriguez *et al.*, 2007).

De plus, il est intéressant de noter que les trois protéines PU.1, c-Myb et CEBP β ciblées par les miARN gga-miR-155 et mdv1-miR-M4-5P semblent coopérer dans leurs fonctions transactivatrices. En effet, ces protéines correspondent aux principaux transactivateurs de différentes protéases à sérine de neutrophiles (Cathepsin G, Neutrophile elastase, protéinase 3 et la myeloblastine) (Lennartsson *et al.*, 2005; Lutz *et al.*, 2001; Oelgeschlager *et al.*, 1996), qui sont des protéines régulatrices importantes dans la réponse immunitaire inflammatoire innée et activent spécifiquement les cytokines pro-inflammatoires telles qu'IL-1 β et TNF α (Meyer-Hoffert, 2009). Les facteurs c-Myb et CEBP β sont également capables de former un complexe caractérisé comme un régulateur clef de différents gènes spécifiques des cellules myéloïdes tels que tom-1A et mim-1 (Plachetka *et al.*, 2008; Tahirov *et al.*, 2002). Ainsi, la diminution de ces facteurs par miR-155 ou un de ces orthologues pourrait induire une répression indirecte potentiellement synergique de ces différents gènes impliqués dans la différenciation cellulaire, le développement, la prolifération ou la réponse immunitaire. De plus, toutes ces données montrent la potentielle importance de mdv1-miR-M4 dans la réponse immunitaire et concordent avec la littérature décrivant une forte diminution du taux de nombreuses interleukines au cours de l'infection par GaHV-2 (Heidari *et al.*, 2008).

En plus de son potentiel rôle dans l'inhibition de la réponse immunitaire cellulaire en tant qu'orthologue de gga-miR-155, il a été montré que mdv1-miR-M4 était également impliqué dans la lymphomagenèse induite par GaHV-2. En effet, grâce à l'utilisation de bacmides recombinants délétés des deux copies de mdv1-miR-M4 ou simplement mutés pour deux nucléotides de sa séquence « seed », Zhao et ses collaborateurs ont montré une très forte diminution de l'apparition de tumeurs et de l'incidence de GaHV-2 sur la survie de poulets B¹⁹/B¹⁹ (Zhao *et al.*, 2011). Le mécanisme précis

fonctions de la cible	Gènes cibles	orthologue
Facteurs de transcription	ARID2 (AT rich interactive domain 2)	
	BACH1 (BTB and CNC homology 1, basic leucine zipper transcription factor1)	KSHV
	CEBPβ (CCAAT/enhancer binding protein, beta)	GaHV-2
	c-Myb (v-myb myeloblastosis viral oncogene homolog (avian))	GaHV-2
	ETS1 (v-ets erythroblastosis virus E26 oncogene homolog 1)	
	FOS v-fos FBJ murine osteosarcoma viral oncogene homolog	KSHV
	FOXO3a (Forkhead box O3A)	
	HIF (Hypoxia-inducible factor 1)	
	HIVEP2 (Human immunodeficiency virus type I enhancer binding protein 2)	GaHV-2, KSHV
	MAF (v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian))	
	MEIS1 (Meis homeobox 1)	
	PU.1 ((SPI1) Spleen focus forming virus (SFFV) proviral integration oncogene spi1)	GaHV-2
	RREB1 (Ras responsive binding protein1)	GaHV-2
	RUNX2 (runt-related transcription factor2)	
	SMAD1 (SMAD family member 1)	
	SMAD2 (SMAD family member 2)	
	SMAD5 (SMAD family member 5)	
	ZIC3 (Zic family member 3)	
	ZNF652 (Zinc finger protein 652)	
Recepteurs	AT1R (Angiotensin II receptor, type 1)	
	AGTRAP angiotensin II receptor-associated protein	KSHV
	CSF1R (Colony stimulating factor 1 receptor)	
	IL13Rα1 (IL13 receptor α1)	
Kinases	IKKε (Inhibitor of kappa light polypeptide gene enhancer in B cells, kinase ε)	KSHV
	MAP3K7IP2 (ou TAB2 , Mitogen-activated protein kinase kinase kinase7 interacting protein 2)	GaHV-2
	MAP3K10 (mitogen-activated protein kinase kinase kinase 10)	KSHV
	PIK3CA (phosphoinositide-3-kinase,catalytic, alpha polypeptide)	KSHV
	RIPK1 (Receptor (TNFRSF)-interacting serine–threonine kinase 1)	
	RFK (Riboflavin kinase)	KSHV
	RPS6KA3 (ribosomal protein S6 kinase,90kDa, polypeptide 3)	
	SOCS1 (suppressor of cytokine signaling 1)	
	TWF1 twinfilin, actin-binding protein, homolog1 (drosophila)	KSHV
Protéines nucléaires	AID (Activation-induced cytidine deaminase)	
	ARNTL (Aryl hydrocarbon receptor nuclear translocator-like)	
	CUTL1 ((CUX1) Cut-like homeobox 1)	
	JARID2 (Jumonji, AT rich interactive domain 2)	
	TP53INP1 (Tumor protein p53 induced nuclear protein 1)	
Protéines adaptatrices	FADD (Fas (TNFRSF6)-associated via death domain protein)	
	MLSTD2 (male sterility domain containing 2)	KSHV
	PICALM (Phosphatidylinositol binding clathrin assembly protein)	
	RHOA (Ras homolog gene family, member A)	
	SLA (Scr-like-adaptor)	KSHV
	MYD88 (myeloid differentiation primary response gene (88))	
Autres	BCL2L13 (BCL2-like 13 (apoptosis facilitator))	GaHV-2
	PDCD6 (Programme cell death 6)	GaHV-2
	BIRC4BP (XIAP associated factor-1)	KSHV
	INPP5D (inositol polyphosphate-5-phosphatase)	
	KGF (Keratynocyte growth factor)	
	MYO10 (myosinX)	
	MSH2 (mutS homolog 2)	
	MSH6 (mutS homolog 6)	
	MLH1 (mutL homolog 1)	
	SAMHD1 (SAM domain and HD domain 1)	KSHV

Tableau 26 : Cibles validées pour miR-155 et/ou ses orthologues viraux identifiés chez HHV-8 (KSHV) et GaHV-2.

par lequel mdv1-miR-M4 est impliqué dans l'oncogenèse induite par GaHV-2 n'est pas encore établi ; cependant au vu des différentes cibles identifiées pour miR-155 ou ses orthologues (Tableau 26), de nombreuses voies cellulaires impliquées dans la lymphomagenèse sont touchées comme par exemple, le contrôle des gènes proapoptotiques *BCL2L13* et *PDCD6*, ciblés par mdv1-miR-M4-5P (Zhao *et al.*, 2009b). De même un rôle potentiel dans la lymphomagenèse a été suspecté pour un autre orthologue du miR-155 cellulaire, le kshv-miR-K12-11 retrouvé au sein du panel de miARN viraux exprimés lors de l'infection par le virus lymphotrope HHV-8 (Gottwein *et al.*, 2007; Skalsky *et al.*, 2007). Cependant, mdv1-pré-miR-M4 semble plus proche de son orthologue aviaire gga-pré-miR-155 (homologies significatives pour le miR-5P et la boucle de gga-miR-155), que ne l'est l'orthologue kshv-miR-K12-11 du miARN humain hsa-miR-155. En effet, pour ces deux pré-miARN, les deux miARN orthologues correspondent respectivement au miR-3P et au miR-5P (Figure 71). Il est intéressant de noter que chez HHV-6 dans lequel aucun orthologue de miR-155 n'a été identifié, la protéine LMP1 transactive, durant la phase de latence III, l'expression du hsa-miR-155 par les voies NF- κ B et p38/MAPK (Rahadiani *et al.*, 2008). De plus, il a été proposé dans ce modèle que hsa-miR-155 serait impliqué dans la réactivation du virus en ciblant les protéines de la cascade de signalisation BMP (Bone Morphogenetic protein) qui possèdent, entre autres, un effet antitumoral (Yin *et al.*, 2010). Ainsi, bien qu'un effet direct sur l'oncogenèse n'ait été démontré que pour GaHV-2, il semblerait que le miR-155 ou ses orthologues joueraient un rôle important dans le déroulement du cycle viral des herpesvirus et dans l'établissement de la lymphomagenèse. A ce jour, plus d'une cinquantaine de cibles ont été validées pour miR-155 ou ses orthologues (Tableau 26) et les mécanismes précis de l'implication et du rôle de miR-155 ou de ces orthologues au cours des infections virales ne sont pas encore totalement caractérisés. Enfin, l'impact de miR-155 ne semble pas restreint au monde des herpesvirus puisqu'une surexpression de ce dernier a également été décrite lors d'infection par le virus HTLV-1 (Human T-cell leukemia virus type-I) (Bellon *et al.*, 2009), ou par le REV-T (Reticuloendotheliosis virus strain T) (Bolisetty *et al.*, 2009).

Dans le cadre de la recherche d'une implication potentielle de gga-miR-155 ou de son orthologue dans le déroulement du cycle viral, une cible potentielle pour les deux miARN de mdv1-pré-miR-M4 (M4-5P et M4-3P) a été mise en évidence respectivement au sein des séquences codantes des gènes UL-28 et UL-32 de GaHV-2. Il faut noter que l'existence de miRE fonctionnels au sein des ORF, largement admise chez la drosophile, est actuellement de plus en plus documentée chez les mammifères (Fang & Rajewsky, 2011; Gu *et al.*, 2009; Moretti *et al.*, 2010; Schnall-Levin *et al.*, 2010; Shin *et al.*, 2010; Tay *et al.*, 2008). Ainsi, le gène UL-28 présentait une complémentarité avec les nucléotides 2-12 de mdv1-miR-M4-5P qui est le miARN le plus exprimé durant l'infection par GaHV-2. Tandis qu'UL-32 présentait une complémentarité avec les nucléotides 1-12 de mdv1-miR-M4-3P qui a été retrouvé faiblement exprimé (autour de 0,5% des miARN viraux) *in vivo* dans les PBL issus de poulets infectés par la souche GaHV-2 RB-1B ou par la souche vaccinale GaHV-2 Rispens et dans les deux lignées lymphoblastoïdes MSB-1 et 54-O issues de tumeur induite par

GaHV-2. Nous avons montré, à l'aide d'expression ectopique de mdv1-miR-M4 (Figure 53) et en système rapporteur luciférase dans des cellules exprimant de façon constitutive ce miARN (Figure 52A), que le mdv1-miR-M4-5P inhibait l'expression de la protéine UL-28 alors que l'orthologue cellulaire gga-miR-155 n'avait aucun effet sur l'expression de ce gène. Cette différence d'action entre miR-155 et son orthologue mdv1-miR-M4-5P pourrait s'expliquer par la complémentarité plus élevée entre le miARN viral et la cible (complémentarité 2-12) en comparaison d'une complémentarité restreinte à la séquence « seed » (nt 2-8) observée entre gga-miR-155 et UL-28 (Figure 51B). De plus lors de ces essais, il est apparu que le miARN mdv1-miR-M4-3P, bien que moins exprimé, réprimait efficacement l'expression de la protéine UL-32 (Figure 53). Ces hypothèses sont confortées par des essais récents fondés sur l'utilisation de bacmides recombinants délétés pour mdv1-miR-M4 (Zhao *et al.*, 2011). En effet, si ces auteurs ont montré que la délétion de mdv1-miR-M4 ne semblait pas avoir d'impact sur le cycle lytique du virus en cellules primaires CEF, le remplacement de mdv1-miR-M4 par gga-miR-155 induisait, en revanche, une apparition plus tardive des tumeurs induites par GaHV-2 suggérant un rôle important des nucléotides localisés en dehors de la séquence « seed » de ces deux miARN et/ou de la présence de mdv1-miR-M4-3P (Zhao *et al.*, 2011).

Il faut noter que les protéines UL28 et UL32 ne sont pas caractérisées chez GaHV-2, alors que les protéines homologues chez HHV-1 ont été identifiées comme nécessaires au clivage et à l'encapsidation de l'ADN viral au sein du virion. De plus, le fait que ces protéines seraient codées par des gènes « core » fortement conservés au sein des herpesvirus de mammifères et aviaires, semble permettre de supposer que bien que non encore démontré dans notre modèle, ces protéines seraient impliquées dans le mécanisme d'empaquetage de l'ADN viral au sein des virions de GaHV-2. Dans la mesure où ces protéines UL-28 et UL-32 seraient impliquées dans le stade tardif de la réplication de GaHV-2, la diminution de leur expression par mdv1-miR-M4 participerait au maintien de la latence. Par ailleurs, il est intéressant de noter qu'une régulation de l'expression de la protéine UL-32 pourrait également être suspectée chez HHV-1. En effet le miARN hsv1-miR-H16 de HHV-1 serait transcrit à partir du brin complémentaire de la région codante du gène *UL-32* (Jurak *et al.*, 2010). Bien que son impact sur la production de UL-32 n'ait pas été étudié, la position du miARN laisse à penser à une potentielle régulation de cette protéine par un mécanisme de type siARN du fait de la complémentarité parfaite entre ce miARN et sa cible.

Pour terminer sur cette partie concernant la régulation des gènes viraux par les miARN viraux, chez les herpesvirus, plusieurs miARN viraux ont été identifiés comme interférant avec le cycle viral. En effet, des gènes très précoces tels que ICP0, ICP4, ICP34.5 et IE72 sont des cibles de miARN viraux chez HHV-1, HHV-2 et HHV-5 (Murphy *et al.*, 2008; Tang *et al.*, 2008; Tang *et al.*, 2009; Umbach *et al.*, 2008). De plus, chez HHV-6, il a été décrit un contrôle du cycle viral via le ciblage de la polymérase virale par des miARN (Barth *et al.*, 2008). Par ailleurs, il faut noter que ce mécanisme ne semble pas exclusif des herpesvirus puisqu'un mécanisme similaire a été identifié chez le virus d'insecte HvAV (*Heliothis virescens* ascovirus) (Hussain *et al.*, 2008).

A l'image d'HHV-1 et de la plupart des *alpha-herpesvirinae*, le génome de GaHV-2 possède des régions répétées inversées IR_L/TR_L et IR_S/TR_S codant pour la majorité de leurs miARN. Cependant, l'organisation interne des régions R_L et R_S diffère entre ces deux virus. En effet, contrairement aux régions d'HHV-1 montrant une continuité transcriptionnelle entre les régions R_L et R_S avec la présence des transcrits LAT recouvrant les deux régions, les régions R_L et R_S de GaHV-2 semblent former deux blocs bien distincts d'autant qu'elles sont séparées par une séquence télomérique longue d'environ 800 nt. De plus, la région R_S de GaHV-2, semble, à elle seule, correspondre à la somme des deux régions R_L et R_S de HHV-1. En effet, pour HHV-1 les transcrits LAT et le gène ICP4 qui ne sont pas chevauchants, recouvrent la quasi-totalité des deux régions IR_L/IR_S. Plus précisément, le grand transcrit LAT s'étend sur la majorité de la région IR_L et le premier cinquième de la région IR_S, les 4/5^{ème} restants codant pour le gène *ICP4*. Par contre, chez GaHV-2, les *LAT* et le gène *ICP4* (seule ICP identifiée au sein des régions R_L et R_S de GaHV-2) sont chevauchants et tous deux localisés dans les régions R_S. De plus, contrairement à HHV-1 et aux autres α -herpesvirus, les régions R_L de GaHV-2 contiennent plusieurs gènes impliqués dans l'oncogenèse tels que *meq* (Liu *et al.*, 1998; Liu *et al.*, 1999b), *vTR* (Djeraba-AitLounis *et al.*, 2004; Trapp *et al.*, 2006) ou les miARN (Zhao *et al.*, 2011), les rapprochant en cela des γ -herpesvirus.

Ainsi, cette région IR_L/TR_L de GaHV-2, codant pour un grand nombre d'ORF pour la plupart validées, présente à l'instar des régions R_S de GaHV-2 ou des régions R_L et R_S de HHV-1, des profils transcriptionnels complexes. En effet, contrairement aux gènes présents dans les régions « core » uniques U_S et U_L qui excepté pour la lipase virale et la terminase (U_L15) ne semblent pas subir d'épissage, des phénomènes d'épissage différentiel complexe sont observés dans les régions R_L et R_S de GaHV-2. Aussi, nos travaux plus exhaustifs sur la transcription de la région IR_L/TR_L ont complété des résultats sommaires décrivant un épissage plutôt conventionnel pour les gènes de la famille des protéines de 14kD permettant la formation de deux protéines possédant une extrémité C-terminale identique (Peng *et al.*, 1992), mais également un épissage plus complexe entre le gène *vIL-8* et les gènes *meq*, *R-LORF5a* et *R-LORF4* (Figure 72A) (Jarosinski & Schat, 2006). Ainsi, nous avons identifié de nouveaux transcrits alternatifs correspondant aux régions déjà décrites comme épissés, mais également de grands transcrits alternatifs recouvrant la quasi-totalité de la région et chevauchant en partie les exons précédemment décrits dans la littérature. Plus précisément, ces grands transcrits alternatifs recouvrent environ 10kpb de la région IR_L/TR_L. Ils semblent s'initier au niveau des promoteurs des gènes codant les protéines de 14kD et s'épissent avec l'exon *meq* et les exons de *vIL-8*. De plus, à l'instar de HHV-8, chez lequel la région codant les miARN et les oncogènes viraux est transcrite en plusieurs ARNm épissés à partir de plusieurs promoteurs spécifiques de la phase lytique ou de la latence, ces différents transcrits alternatifs sont initiés à partir de quatre promoteurs également alternatifs, impliqués dans la latence et/ou le cycle lytique de GaHV-2. En effet, en plus du promoteur plutôt spécifique de la latence p-miR-M9-M4 localisé dans cette étude en amont des miR-M9 à M4, permettant la transcription de ces derniers en un messenger et également la synthèse de transcrits

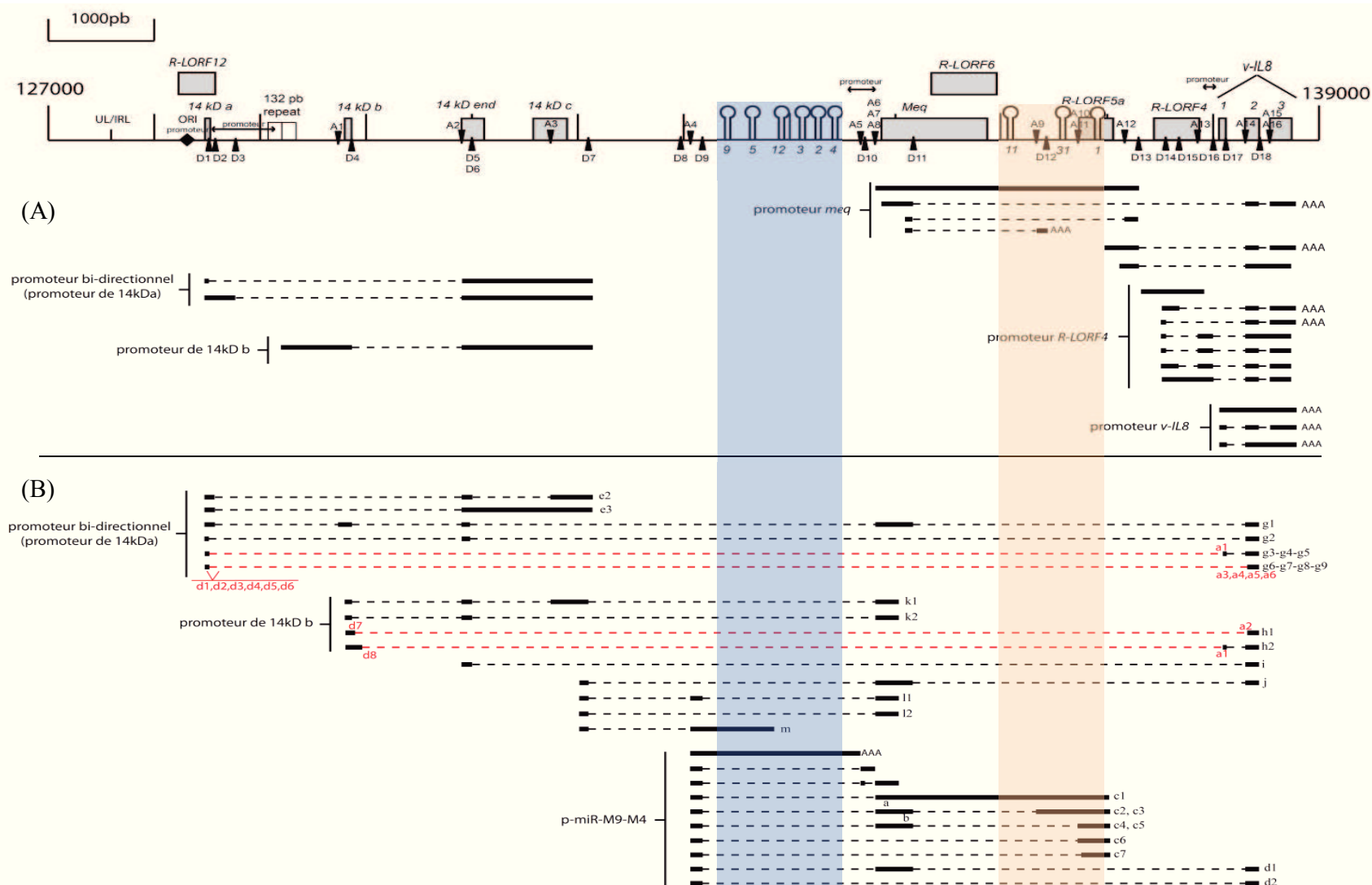


Figure 72 : Carte transcriptionnelle du brin codant les miARN de la région IR_L/IR_S de GaHV-2.

La région comprise entre les bases 127000 et 139000 de GaHV-2 Md5 est représentée dans la partie supérieure, les gènes sont représentés par des rectangles gris et les miARN par des tiges boucles. Les bornes d'épissage accepteuses sont schématisées par des triangles noirs inversés et les bornes d'épissage donneuses par des triangles noirs. (A) Représentation des transcrits identifiés dans la littérature. (B) Représentation des transcrits identifiés au cours de notre étude. Les exons sont représentés par des traits plein noirs, les introns de type U2 par des pointillés noirs et les introns de type U12 par des pointillés rouges. Les transcrits possédant les mêmes exons mais possédant une borne d'épissage alternative proche ne sont figurés que sous la forme d'un seul transcrit. Les queues polyA sont représentées par AAA et les promoteurs potentiels des transcrits sont indiqués à gauche de ces derniers.

alternatifs se terminant en aval du gène *vIL-8*, quatre promoteurs avaient été identifiés précédemment dans la littérature (Figure 72A). Le promoteur le plus en amont de la région R_L a été relativement bien caractérisé. Il correspond au promoteur bidirectionnel localisé au niveau de l'origine de réplication de GaHV-2, permettant la transcription des gènes *pp38/pp24* et des gènes des protéines 14kDa et 14kDc (Parcells *et al.*, 2003). Le deuxième promoteur, localisé entre l'exon 14kDa et les répétitions de 132pb (Tahiri-Alaoui *et al.*, 2009a) permet la transcription des protéines 14kDb et 14kDc. En effet, les transcrits issus de ce promoteur contiennent deux IRES, un premier localisé entre les répétitions de 132pb et 14kDb favorisant la traduction de ce dernier (Tahiri-Alaoui *et al.*, 2009a), et un second également identifié dans les transcrits encodant 14kDa, qui est localisé en aval de l'exon 14kDend et permet la traduction de 14kDc (Tahiri-Alaoui *et al.*, 2009b). Par contre, nous n'avons pas retrouvé le promoteur localisé juste en amont de l'oncogène *meq* et permettant la transcription de ce dernier dans les cellules RP1 transformées par une souche de GaHV-2 de virulence moindre, qui avait été précédemment décrit dans la littérature (Jones *et al.*, 1992). Dans notre modèle, l'intégralité des initiations identifiées pour les transcrits du gène *meq* étaient localisées en amont des miARN au sein de la région promotrice p-miR-M9-M4 (Figure 72B). Enfin, le quatrième promoteur est situé en amont du gène *vIL-8* et permet la transcription de *vIL-8* durant la phase lytique de GaHV-2. L'équipe de Jarosinski avait précédemment remis en question l'importance de l'exon 1 de ce gène composé de trois exons, suite à la découverte de nombreux transcrits alternatifs entre le gène *RLORF4* et les exons 2 et 3 de *vIL-8* (Jarosinski & Schat, 2006). En effet, ces deux derniers exons de *vIL-8* encodent pour les quatre cystéines dont le positionnement est conservé dans toutes les chémokines de type CXC ainsi que trois résidus DKR (aspartate-leucine-arginine), localisés juste en amont des motifs CXC nécessaires à la fonctionnalité de *vIL-8*. Ces 3 résidus sont cependant différents de ceux (ELR ; glutamate-lysine-arginine) classiquement décrits comme essentiels pour la fonction de ces chémokines de type CXC cellulaires. Enfin, il semblerait que l'exon 1 correspondrait plutôt à un peptide signal du fait de sa composition riche en résidus hydrophobes (Parcells *et al.*, 2001). De nouveaux transcrits identifiés au cours de notre étude s'épissaient également avec le gène *vIL-8*. Cependant, la recherche de cadres ouverts de lectures au sein de ces transcrits induisant des variants de *vIL-8* révèlent qu'en dehors de l'épissage avec l'exon *meq*, ces transcrits ne sont pas en phase avec les exons 2 et 3 de *vIL-8*. Ainsi, le rôle de ces transcrits ne serait pas d'augmenter la production de protéines homologues à *vIL-8*. Ils pourraient simplement servir à la transcription des miARN de la région IR_L/TR_L . Par ailleurs, en cellules MSB-1, 78% des transcrits s'initiaient au sein de la région promotrice p-miR-M9-M4 permettant notamment la transcription de *Meq/vIL-8*, tandis que 22% s'initiaient au sein du promoteur de *vIL-8* (Figure 60). A l'inverse, lors de l'infection lytique de GaHV-2 en CEF, 74% des transcrits s'initiaient au sein du promoteur de *vIL-8* déjà identifié à partir de cellules MSB-1 réactivées par du *n*-butyrate (Parcells *et al.*, 2001), et seulement 4% s'initiaient au sein de la région promotrice p-miR-M9-M4. Ainsi, ces données montrent que le promoteur identifié en amont des miARN, responsable de la transcription de l'ensemble des miARN de la région IR_L/TR_L , de *meq* et de ses

transcrits alternatifs serait plutôt un promoteur spécifique de la phase de latence/oncogenèse de GaHV-2. De plus, 20% des initiations obtenues lors de l'infection lytique de GaHV-2 en CEF étaient localisées au début de l'exon 2 de *vIL-8* (Figure 60) et semblaient conduire à des transcrits capables d'encoder seulement une protéine potentielle de 32 AA correspondant à l'extrémité C-terminale de la protéine *vIL-8*, laissant penser à un autre rôle encore inconnu du gène *vIL-8*.

Enfin, la caractérisation du promoteur p-miR-M9-M4 reste encore à être améliorée. Cependant, grâce à l'approche par troncature du promoteur initial défini arbitrairement comme la région entre le gène codant la 14kDc et *mdv1-miR-M5*, nous avons montré que la région minimale permettant l'activité promotrice maximale (80% de l'activité du promoteur initial) correspondait au promoteur Trq-2 (Figure 63). Par ailleurs, parmi les ER identifiées par bioinformatique (Genomatix) dans la partie 5' de Trq-2 absente de la construction Trq3, dont l'activité promotrice n'est qu'à 50% de celle du promoteur initial, trois sites AP-1 potentiellement capables de fixer l'hétérodimère Meq/Jun ont retenu notre attention. Il serait donc intéressant de muter ces trois sites afin de vérifier leur implication dans l'activité promotrice, d'autant que Zhao et ses collaborateurs semblent avoir identifié un rôle de Meq dans la production des miARN du cluster M9-M4 (Zhao *et al.*, 2011). Par ailleurs, ce promoteur semble être le promoteur dominant responsable de la transcription de miARN au cours de la phase de latence/oncogenèse induite par GaHV-2. En effet, les RACE-PCR réalisées à partir de cellules MSB-1 au niveau du cluster M9-M4 présentaient des initiations au niveau de cette région promotrice dans la totalité des clones.

Il faut noter que chez les vertébrés, les pri-miARN semblent transcrits essentiellement à partir de promoteurs spécifiques. En effet, d'une part, la moitié des pri-miARN des vertébrés sont transcrits sous la forme d'un ARNm capé et polyadénylé et d'autre part, parmi l'autre moitié transcrite à partir d'introns de gènes, 35% possèderaient également leur propre promoteur qui peut être localisé très en amont du pri-miARN (Davis & Hata, 2009; Monteys *et al.*, 2010). Les herpesvirus qui sont à ce jour la seule famille de virus codant pour de nombreux miARN au sein de leur génome présentent à une échelle moindre tous les mécanismes moléculaires recensés dans les cellules eucaryotes, notamment ceux concernant la transcription des miARN. De plus, actuellement, certains résultats obtenus dans le modèle herpesvirus ont été à l'origine d'études de mécanismes cellulaires. Toutefois, du fait de la promiscuité des gènes et d'un fort taux d'épissage alternatif, la régulation des miARN viraux apparaît très variée et souvent complexe. A titre d'exemple, l'étude du miR-155 a pris un essor considérable suite à la découverte des orthologues viraux. De même, la transcription des miARN du MHV-68 (murine gammaherpesvirus type 68) localisés au sein du cluster polycistronique d'ARNt a été montrée dépendante de l'ARN pol III (Pfeffer *et al.*, 2005) plusieurs années avant qu'un mécanisme similaire n'ait été identifié dans le modèle murin (Reese *et al.*, 2010). Cependant, au sein des herpesvirus, le processus le plus courant conduisant à la transcription de miARN semble être l'élaboration à partir d'exons ou d'introns de grands ARN non codants associés à la latence, tels que le gène *BART* chez HHV-4 (Edwards *et al.*, 2008; Pfeffer *et al.*, 2004), ou les gènes *LAT* chez HHV-1 (Umbach *et al.*,

2008), HHV-2 (Tang *et al.*, 2008) ou GaHV-2 (Burnside *et al.*, 2006). De même, les miARN de la région BHRF1 chez HHV-4 (Cai *et al.*, 2006; Xing & Kieff, 2007) sont localisés dans les introns du transcrit codant l'antigène nucléaire EBNA1 (Epstein Barr virus Nuclear Antigen), et la majorité des miARN de HHV-8 sont localisés dans de grands introns des transcrits codant pour les formes de la kaposine (K12) au sein de la région associée à la latence contenant également les protéines LANA, v-cyclin, vFLIP (Cai & Cullen, 2006). Enfin, l'étude de l'expression des miARN introniques a diminué la grande part d'ombre existant dans la compréhension des mécanismes précis régissant la maturation des ARN. En effet, à titre d'exemple, chez HHV-4, il a été montré pour 14 miARN de BART, que les miARN étaient excisés par la machinerie cellulaire (Drosha) avant même l'épissage du gène (Edwards *et al.*, 2008).

Pour finir, chez les herpesvirus humains HHV-1, HHV-2 et HHV-8, des miARN ont également été décrits au sein de régions codant pour des protéines. En effet, chez HHV-8, deux miARN sont localisés dans un exon commun aux différents transcrits codant les kaposines (Cai & Cullen, 2006), tandis que chez HHV-1 et HHV-2, des miARN sont identifiés dans les régions codantes des gènes ICP0, ICP4 et ICP34.5 (Jurak *et al.*, 2010).

La comparaison de la transcription des miARN de l'IR_L/TR_L de GaHV-2 avec les différents systèmes de transcription des miARN chez les herpesvirus, semble la rapprocher de celle observée pour les miR-K12-1 à miR-K12-12 de HHV-8. En effet, la plupart des miARN (cluster miR-1 à 9,11) de HHV-8 sont transcrits au sein de deux types de grands introns d'environ cinq et neuf kpb à partir de trois promoteurs plutôt spécifiques de la latence (Figure 73). Par ailleurs, les deux autres miARN (kshv-miR-K12-10 et 12) localisés dans l'exon contenant le gène de la kaposine sont exprimés à partir du même transcrit pré-ARNm que le cluster intronique miR-K12-1 à 9,11. De plus, pour ces deux miARN, un autre promoteur spécifique du gène de la kaposine exprimé pendant la phase lytique du virus permettrait leur transcription au sein d'un transcrit non épissé (Figure 73) (Cai & Cullen, 2006). De la même manière, dans notre modèle viral GaHV-2 nous avons montré que les miARN de la région IRL /TRL pouvaient résulter de mécanismes de transcription complexes, dont l'utilisation de différents promoteurs, la présence de transcrits non épissés ou épissés et une localisation au sein de ces derniers dans des introns ou des exons (Figure 73). En effet, nous avons tout d'abord montré, par RACE-PCR et RT-PCR à partir d'ARN de cellules MSB-1 dans lesquelles le virus GaHV2 est maintenu en latence, la présence d'un ARNm cappé et polyadénylé non épissé comportant exclusivement les miARN mdv1-miR-M9 à M4 (Figure 54 et 55). De plus, ce messenger correspondant au pri-miR-M9-M4 serait sous la dépendance d'un promoteur spécifique de la latence plutôt que de la phase répllicative. De même, la transcription de ce cluster de miARN n'apparaît pas aussi simple que la seule formation de ce transcrit non épissé. En effet, l'ensemble des RT-PCR réalisées à l'aide d'amorces situées en dehors de ce transcrit fait ressortir le cluster M9-M4 sous la forme d'un seul intron. Plus précisément, ces RT-PCR ont permis d'identifier 21 formes d'introns différents issus de multiples bornes donneuses et acceptuses alternatives. Ces grands introns peuvent

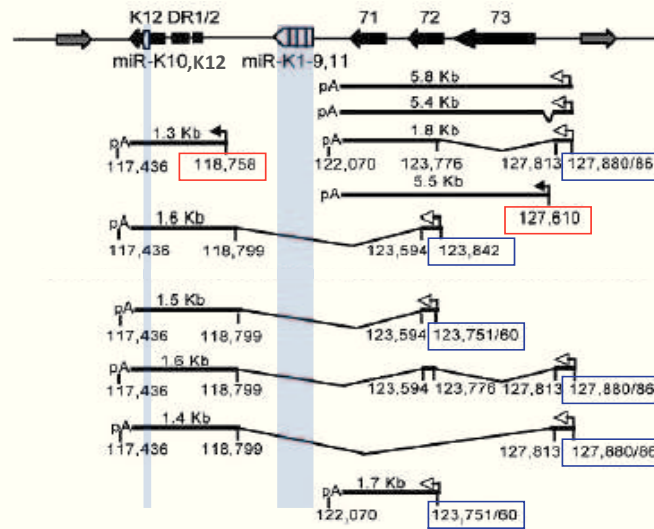


Figure 73 : Profil d'épissage dans la région des miARN du virus de Kaposi (D'après Cai *et al.*, 2006).

Le cluster de miARN miR-K1-9,11 est identifié par la flèche hachurée. Les ORF K12, 71, 72 et 73 codant respectivement pour les protéines K12 (Kaposine), v-FLIP, v-cyclin et LANA, sont représentés par une flèche pleine. Les différents ARNm épissés de la région sont représentés en dessous. Les promoteurs spécifiques de la latence sont indiqués par une flèche blanche et un rectangle bleu et ceux actifs durant la phase lytique, par une flèche noire et un rectangle rouge. La localisation des bornes d'épissage, des promoteurs et des polyA est indiquée sous chaque ARNm.

être rapprochés de ceux observés chez HHV-8 (Cai & Cullen, 2006), mais également de ceux identifiés au sein du gène *LAT* de GaHV-2 qui présente de nombreux ARNm hautement épissés et polyadénylés (Cantello *et al.*, 1994; Cantello *et al.*, 1997; Li *et al.*, 1998; Li *et al.*, 1994); communication personnelle. S. Laurent)).

Jusqu'ici les miARN de cette région étaient identifiés dans la littérature comme un seul et unique cluster s'étendant du miR-M9 au miR-M1 sur une longueur de plus de 3500 nucléotides et renfermant entre autre l'oncogène *meq*. Les résultats précédents, identifiant le pré-miR 9-4 dans un ARNm propre et surtout ne montrant jamais un épissage quelconque regroupant les miARN en amont et en aval de *meq* autrement qu'au sein d'un grand intron, sont en faveur de la définition d'un cluster miR-M9-M4 indépendant des trois autres miARN. Par contre, ces miARN mdv1-miR-M11, M31 et M1 non seulement n'appartiennent pas au même cluster que les miR-M9-M4 mais, de plus, leur regroupement au sein d'un même cluster ne semble pas clair. En effet, bien que ces trois miARN soient proches au niveau de leur localisation génomique et qu'ils aient été observés au sein des différents transcrits du gène *meq*, de façon exonique au sein du grand transcrit non épissé recouvrant également le gène *R-LORF5a* ou de façon intronique au sein des transcrits épissés du gène *meq* (Figure 72A) (Peng *et al.*, 1995), différentes données obtenues au cours de notre étude conduisent à douter de leur expression exclusivement sous la forme d'un cluster. En effet, s'ils peuvent tous être transcrits à partir d'un même intron ou d'un même exon, un épissage interne à la région codant ces trois miARN peut également les séparer en deux parties au sein de différents exons et introns (Figure 72B).

Au vu de ces multiples pré-ARNm permettant la transcription des miARN de la région IR_L/TR_L, nous avons voulu tester l'hypothèse d'une corrélation potentielle entre ces événements d'épissage et les différences d'expression observées entre les miARN de GaHV-2 et plus particulièrement, au sein du cluster M9-M4. A titre d'exemple, pour ne donner que les 2 extrêmes, le mdv1-miR-M9, le premier miARN du cluster M9-M4, est détecté en très faible quantité ne représentant qu'environ 0,2% des miARN viraux séquencés à partir de PBL infectés par la souche GaHV-2 RB-1B, alors que le mdv1-miR-M4, le dernier élément de ce même cluster, représente 66% de ces mêmes miARN viraux totaux. Plus globalement, si l'on considère l'ensemble de tous les miARN de la région IR_L/TR_L, six d'entre eux (mdv1-miR-M5, M12, M3, M2, M4 et M1) sont fortement exprimés selon les banques analysées, avec des variations détectables en fonction de la phase du cycle viral et des méthodes de séquençage (Tableau 16). Cependant, mdv1-miR-M4-5P a toujours été retrouvé fortement exprimé, pouvant même représenter à lui seul jusqu'à 72% des miARN de GaHV-2 (Morgan *et al.*, 2008). Au sein du laboratoire, une approche voisine de celle développée dans ma thèse a montré qu'un seul promoteur, piloté par la protéine p53 avait été identifié pour le cluster de miARN (cluster-M8-M10) de la région IR_S/TR_S (Stik *et al.*, 2010) et que la différence d'expression de certains miARN de ce cluster semblait résulter d'un phénomène d'épissage différentiel (communication personnelle P. Rasschaert). Cependant, un épissage différentiel favorisant

la surexpression de mdv1-miR-M4 ne semble pas pouvoir être mis en avant pour le cluster-M9-M4, puisque ce dernier est toujours retrouvé transcrit dans son intégralité, soit sous la forme d'un ARNm correspondant au pri-miARN complet, soit au sein d'introns (Figure 72B). Cependant, on ne peut exclure que les différentes formes d'introns détectés puissent influencer sur la structure du pri-miARN et ainsi favoriser la production de certains miARN.

Sachant que l'épissage ne semble pas expliquer la surexpression de mdv1-miR-M4-5P dans l'ensemble des banques de miARN, une des explications potentielles pouvant rendre compte de cette surexpression est la présence d'un promoteur cryptique propre au mdv1-miR-M4 au sein du cluster M9-M4. Une étude est en cours au sein du laboratoire afin de vérifier cette hypothèse et les résultats préliminaires obtenus montrent une activité promotrice de la région juste en amont de mdv1-miR-M4 bien que le TSS de ce promoteur potentiel n'ait pas encore été déterminé. Nous testons également la fonctionnalité de ce promoteur au moment du cycle lytique et/ou de la latence, car il pourrait rendre compte de la surexpression du mdv1-miR-M4 quelle que soit la phase du cycle viral, qui ne semble pas s'expliquer par le promoteur p-miR-M9-M4 dont l'activité apparaît plutôt spécifique de la phase de latence virale. Enfin, à l'instar d'HHV-4, il se pourrait également que le taux de transcrit primaire n'influence pas nécessairement le niveau des miARN matures et que l'étape clef de la production des miARN soit plutôt la maturation de ces derniers (Amoroso *et al.*, 2011). Par ailleurs, ces dernières données corréleraient avec celles obtenues par Edwards et ses collaborateurs (Edwards *et al.*, 2008), montrant une excision des miARN BART avant les phénomènes d'épissage régulant la région.

Parmi les transcrits permettant la transcription des miARN sous forme intronique, neuf présentaient des introns ne possédant pas de bornes canoniques de type U2 (GT-AG) mais plutôt des bornes de type U12 (AT-AC) (Tableau 20). Peu de données sont retrouvées dans la littérature concernant ce splicéosome minoritaire responsable de l'excision de moins de 0,5% des introns chez l'homme (Burset *et al.*, 2001). Ce mécanisme, conservé au sein des métazoaires, semble être apparu tôt au cours de l'évolution et semble également être à l'origine de l'épissage de type U2 (Bartschat & Samuelsson, 2010; Lin *et al.*, 2010). Ainsi, il semble logique que les virus, qui évoluent en parallèle des cellules eucaryotes, aient également gardé ce système de régulation de l'épissage. Cette description de transcrits potentiellement épissés par le splicéosome de type U12 n'a jamais été décrite chez les virus. Cependant au vu de l'inhibition de l'épissage via des protéines virales, notamment ICP27 dans le cas de GaHV-2 (Amor *et al.*, 2011), il se pourrait que ce mécanisme soit une voie alternative d'épissage lors de l'infection par GaHV-2 permettant l'épissage de la région tandis que d'une manière générale l'épissage de type U2 serait inhibé. Les résultats décrits dans notre travail ont juste permis la description de ces transcrits, et l'implication du splicéosome de type U12 dans la formation de ces transcrits ainsi que son rôle dans la pathogenèse de GaHV-2 restent toutefois à confirmer.

Lors de l'étude des transcrits encodant les miARN de l'IR_L/TR_L de GaHV-2, de nombreux transcrits correspondant au brin inverse des miARN ont été identifiés (Figure 65). Au sein de ces

transcrits nous avons retrouvé un transcrit déjà identifié dans la littérature comme un variant du gène *meq* et dénommé delta-Meq (Okada *et al.*, 2007). Cependant, les données obtenues montrent que ce variant de la protéine Meq, décrit comme un régulateur négatif de Meq et L-Meq est artificiel puisque le transcrit permettant sa formation provient du brin inverse des miARN et de *meq*. Par ailleurs, au vu de l'épissage des gènes présents sur le brin codant les miARN ayant un rôle primordial dans la pathogenèse de GaHV-2, ces transcrits identifiés au cours de notre étude permettent une meilleure connaissance des mécanismes de transcription des gènes du brin inverse des miARN, prédits pour la plupart.

Pour finir, au cours de notre étude un phénomène d'hyper-édition a été identifié sur cinq transcrits issus du brin inverse des miARN ainsi que sur un transcrit issu du brin codant les miARN (Figure 66). Un phénomène d'édition a déjà été décrit chez les herpesvirus HHV-4 (Iizasa *et al.*, 2010) et HHV-8 (Gandy *et al.*, 2007) ; cependant, contrairement à ces deux virus où seule une base apparaissait éditée, de très nombreuses bases (jusqu'à 185 modifications A→I, identifiées comme des mutations A→G) sont retrouvées éditées chez GaHV-2. En effet, ce phénomène d'édition s'étend sur environ 2500 pb de la région IR_L/TR_L et semble toucher jusqu'à la moitié des adénines présentes au niveau des zones éditées. Cette édition des transcrits de GaHV-2 est probablement due à une des protéines ADAR (ADAR1 ou ADAR2), mais plus particulièrement ADAR1L, protéine induite par l'interféron α qui est induit lors de l'infection des cellules par GaHV-2 (Heidari *et al.*, 2008). De plus, le mécanisme observé dans notre étude apparaît non sélectif sur le choix des bases. En effet deux cas de figure sont observés lors de l'édition par les protéines ADAR, un premier dit spécifique lorsque la structure double brin comprend de 15 à 40 nucléotides induisant la modification d'une base particulière et un second, correspondant à ce que l'on a observé pour GaHV-2, induisant la modification de la moitié des adénosines lorsque la structure double brin comprend plus de 50 nucléotides (Hundley & Bass, 2010). Ce phénomène d'hyper-édition a déjà été rapporté chez l'arévirus LCMV, où 38% des adénines des ARNm codant les glycoprotéines sont touchées aboutissant à la traduction de glycoprotéines non fonctionnelles ne pouvant atteindre la lumière du réticulum endoplasmique ou la surface cellulaire, diminuant ainsi le pouvoir infectieux de la population virale (Zahn *et al.*, 2007). Cependant, dans le cas des transcrits édités identifiés au cours de notre étude, la majorité des bases éditées sont retrouvées dans une région non codante et seules les protéines MRLK et R-LORF10 apparaissent touchées au sein des transcrits Edit-185 et Edit-88 (Figure 66).

L'influence de ce phénomène d'édition est multiple et permet notamment la modification de la séquence en AA et l'ouverture de nombreux cadres de lecture (Figure 68) par création de codon initiateurs AUG et gommage des codons STOP (Figure 67). De plus, l'édition de ces transcrits modifie fortement la structure secondaire de l'ARN grâce au nouvel appariement possible C-G à la place du mésappariement CxA, tout en gardant les interactions avec les thymine grâce à l'appariement U-G (Figure 74). L'édition est également capable de modifier l'épissage de la région par altération des

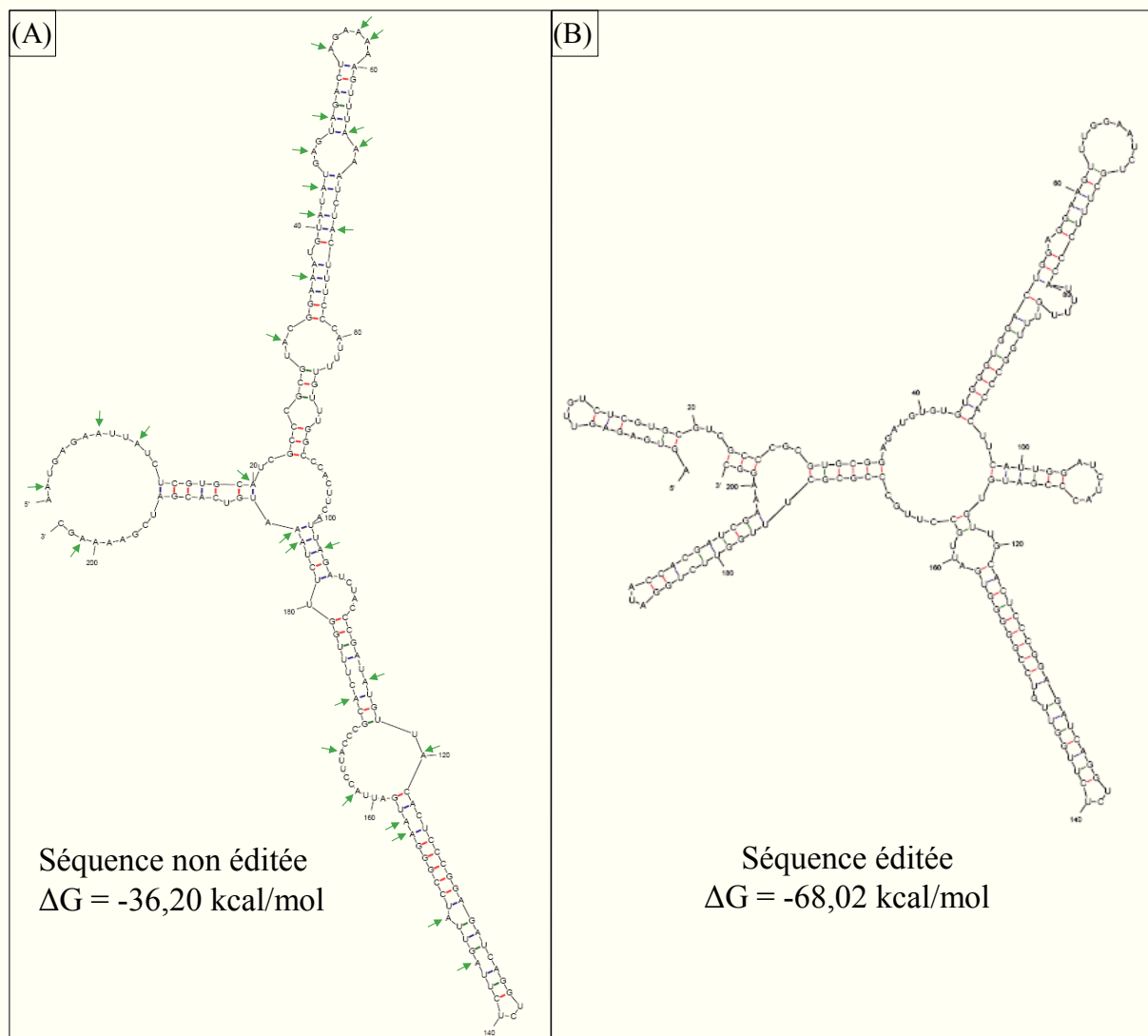


Figure 74 : Prédiction de la structure secondaire de l'ARN de la séquence éditée Edit-29 grâce au logiciel mfold.
 (A) Structure secondaire de l'ARN anti-sens non édité, les flèches vertes indiquent les adénines éditées.
 (B) Structure secondaire de l'ARN anti-sens édité.

bornes donneuses (Edit-185) et accepteuses (Edit-18), bien que ce mécanisme ne semble avoir été décrit jusqu'ici que pour ADAR2 (Samuel, 2011).

Hormis au sein d'ARNm mitochondriaux des trypanosomes où le phénomène d'édition peut modifier 90% des codons du gène *cox3* (subunit 3 of the cytochrome *c* oxydase) par des mécanismes d'insertion et de délétion (Horton & Landweber, 2002), une telle quantité de modifications A→G n'a jamais été identifiée chez les eucaryotes ou les virus. Ainsi, tout comme chez les trypanosomes, la production de protéines ou d'ARN fonctionnels suite à ce phénomène d'hyper-édition n'est pas à exclure. Par ailleurs, la présence d'édition à 140 nucléotides en amont des miARN, observée pour le transcrit Edit-18 pourrait potentiellement toucher les pré-miARN et ainsi modifier la séquence des miARN matures ou inhiber leur maturation.

A l'instar d'HHV-8 montrant une augmentation de l'édition lors de la phase lytique du virus (Gandy *et al.*, 2007), l'étude actuellement en cours au laboratoire a permis de montrer un très fort taux d'édition lors d'une infection lytique en CEF par la souche GaHV-2 RB-1B. En effet, en comparaison des cellules MSB-1 où seuls quelques transcrits édités ont été identifiés, un taux de 60% de séquences éditées a été retrouvé lors de l'infection lytique. Par ailleurs, ces données semblent corrélérer avec le taux d'interféron observé *in vivo* lors de l'infection de poulet par GaHV-2 montrant une augmentation du taux de d'interféron lors de la première phase cytotytique puis une diminution de ce taux lors de la phase de latence/oncogenèse (Heidari *et al.*, 2008).

RÉFÉRENCES BIBLIOGRAPHIQUES

- Abdul-Careem, M. F., Hunter, B. D., Lee, L. F., Fairbrother, J. H., Haghighi, H. R., Read, L., Parvizi, P., Heidari, M. & Sharif, S. (2008). Host responses in the bursa of Fabricius of chickens infected with virulent Marek's disease virus. *Virology* **379**, 256-265.
- Advani, S. J., Brandimarti, R., Weichselbaum, R. R. & Roizman, B. (2000a). The disappearance of cyclins A and B and the increase in activity of the G(2)/M-phase cellular kinase cdc2 in herpes simplex virus 1-infected cells require expression of the alpha22/U(S)1.5 and U(L)13 viral genes. *J Virol* **74**, 8-15.
- Advani, S. J., Weichselbaum, R. R. & Roizman, B. (2000b). The role of cdc2 in the expression of herpes simplex virus genes. *Proc Natl Acad Sci U S A* **97**, 10996-11001.
- Afonso, C. L., Tulman, E. R., Lu, Z., Zsak, L., Rock, D. L. & Kutish, G. F. (2001). The genome of turkey herpesvirus. *J Virol* **75**, 971-978.
- Akira, S., Ishihara, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., Nakajima, T., Hirano, T. & Kishimoto, T. (1990). A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO J* **9**, 1897-1906.
- Akiyama, Y. & Kato, S. (1974). Two cell lines from lymphomas of Marek's disease. *Biken J* **17**, 105-116.
- Alcami, A. & Koszinowski, U. H. (2000). Viral mechanisms of immune evasion. *Trends Microbiol* **8**, 410-418.
- Alfonzo, J. D., Blanc, V., Estevez, A. M., Rubio, M. A. & Simpson, L. (1999). C to U editing of the anticodon of imported mitochondrial tRNA(Trp) allows decoding of the UGA stop codon in *Leishmania tarentolae*. *EMBO J* **18**, 7056-7062.
- Alioto, T. S. (2007). U12DB: a database of orthologous U12-type spliceosomal introns. *Nucleic Acids Res* **35**, D110-115.
- Allen, R. D., 3rd, Kim, H. K., Sarafova, S. D. & Siu, G. (2001). Negative regulation of CD4 gene expression by a HES-1-c-Myb complex. *Mol Cell Biol* **21**, 3071-3082.
- Amor, S., Strassheim, S., Dambrine, G., Remy, S., Rasschaert, D. & Laurent, S. (2011). The ICP27 protein of Marek's disease virus interacts with SR proteins and inhibits the splicing of cellular telomerase chTERT and viral vIL8 transcripts. *J Gen Virol*.
- Amoroso, R., Fitzsimmons, L., Thomas, W. A., Kelly, G. L., Rowe, M. & Bell, A. I. (2011). Quantitative studies of Epstein-Barr virus-encoded microRNAs provide novel insights into their regulation. *J Virol* **85**, 996-1010.
- Anderson, A. S., Francesconi, A. & Morgan, R. W. (1992). Complete nucleotide sequence of the Marek's disease virus ICP4 gene. *Virology* **189**, 657-667.
- Anderson, A. S., Parcells, M. S. & Morgan, R. W. (1998). The glycoprotein D (US6) homolog is not essential for oncogenicity or horizontal transmission of Marek's disease virus. *J Virol* **72**, 2548-2553.
- Anobile, J. M., Arumugaswami, V., Downs, D., Czymmek, K., Parcells, M. & Schmidt, C. J. (2006). Nuclear localization and dynamic properties of the Marek's disease virus oncogene products Meq and Meq/vIL8. *J Virol* **80**, 1160-1166.
- Aparicio, O., Razquin, N., Zaratiegui, M., Narvaiza, I. & Fortes, P. (2006). Adenovirus virus-associated RNA is processed to functional interfering RNAs involved in virus production. *J Virol* **80**, 1376-1384.
- Baba, T. W., Giroir, B. P. & Humphries, E. H. (1985). Cell lines derived from avian lymphomas exhibit two distinct phenotypes. *Virology* **144**, 139-151.
- Baigent, S. & Nair, V. (2004). Marek's disease virus: biology and life cycle. In *Marek's Disease: An evolving Problem*, pp. 62-77. Edited by F. Davison & V. Nair. London: Elsevier Academic Press.
- Baker, K. E. & Parker, R. (2004). Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol* **16**, 293-299.
- Barrow, A. D., Burgess, S. C., Baigent, S. J., Howes, K. & Nair, V. K. (2003). Infection of macrophages by a lymphotropic herpesvirus: a new tropism for Marek's disease virus. *J Gen Virol* **84**, 2635-2645.
- Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233.

- Barth, S., Pfuhl, T., Mamiani, A., Ehse, C., Roemer, K., Kremmer, E., Jaker, C., Hock, J., Meister, G. & Grasser, F. A. (2008). Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. *Nucleic Acids Res* **36**, 666-675.
- Bartschat, S. & Samuelsson, T. (2010). U12 type introns were lost at multiple occasions during evolution. *BMC Genomics* **11**, 106.
- Bass, B. L. & Weintraub, H. (1987). A developmentally regulated activity that unwinds RNA duplexes. *Cell* **48**, 607-613.
- Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M. & Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. *Genome Res* **10**, 1001-1010.
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P. & Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev* **20**, 1885-1898.
- Bellon, M., Lepelletier, Y., Hermine, O. & Nicot, C. (2009). Deregulation of microRNA involved in hematopoiesis and the immune response in HTLV-I adult T-cell leukemia. *Blood* **113**, 4914-4917.
- Ben-Shlomo, H., Levitan, A., Shay, N. E., Goncharov, I. & Michaeli, S. (1999). RNA editing associated with the generation of two distinct conformations of the trypanosomatid *Leptomonas collosoma* 7SL RNA. *J Biol Chem* **274**, 25642-25650.
- Bender, T. P., Kremer, C. S., Kraus, M., Buch, T. & Rajewsky, K. (2004). Critical functions for c-Myb at three checkpoints during thymocyte development. *Nat Immunol* **5**, 721-729.
- Benne, R., Van den Burg, J., Brakenhoff, J. P., Sloof, P., Van Boom, J. H. & Tromp, M. C. (1986). Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**, 819-826.
- Benton, W. J. C., M. S (1957). The increased incidence of visceral lymphomatosis in broiler and replacement birds. *Avian Disease*, 320-327.
- Berget, S. M. (1995). Exon recognition in vertebrate splicing. *J Biol Chem* **270**, 2411-2414.
- Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I. & Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* **125**, 1111-1124.
- Bivalkar-Mehla, S., Vakharia, J., Mehla, R., Abreha, M., Kanwar, J. R., Tikoo, A. & Chauhan, A. (2010). Viral RNA silencing suppressors (RSS): Novel strategy of viruses to ablate the host RNA interference (RNAi) defense system. *Virus Res*.
- Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72**, 291-336.
- Blanc V, F. J.-C., Litvak S, Araya A (2002). Réécriture du matériel génétique : fonctions et mécanismes de l'édition de l'ARN. *MEDECINE/SCIENCES* **18**, 181-192.
- Blondeau, C., Marc, D., Courvoisier, K., Vautherot, J. F. & Denesvre, C. (2008). Functional homologies between avian and human alphaherpesvirus VP22 proteins in cell-to-cell spreading as revealed by a new cis-complementation assay. *J Virol* **82**, 9278-9282.
- Bohnsack, M. T., Czaplinski, K. & Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**, 185-191.
- Bolisetty, M. T., Dy, G., Tam, W. & Beemon, K. L. (2009). Reticuloendotheliosis virus strain T induces miR-155, which targets JARID2 and promotes cell survival. *J Virol* **83**, 12009-12017.
- Borchert, G. M., Lanier, W. & Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* **13**, 1097-1101.
- Bradley, G., Hayashi, M., Lancz, G., Tanaka, A. & Nonoyama, M. (1989a). Structure of the Marek's disease virus BamHI-H gene family: genes of putative importance for tumor induction. *J Virol* **63**, 2534-2542.
- Bradley, G., Lancz, G., Tanaka, A. & Nonoyama, M. (1989b). Loss of Marek's disease virus tumorigenicity is associated with truncation of RNAs transcribed within BamHI-H. *J Virol* **63**, 4129-4135.
- Brennecke, J., Stark, A., Russell, R. B. & Cohen, S. M. (2005). Principles of microRNA-target recognition. *PLoS Biol* **3**, e85.
- Brennicke, A., Marchfelder, A. & Binder, S. (1999). RNA editing. *FEMS Microbiol Rev* **23**, 297-316.

- Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L. & Voinnet, O. (2008).** Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **320**, 1185-1190.
- Brown, A. C., Baigent, S. J., Smith, L. P., Chattoo, J. P., Petherbridge, L. J., Hawes, P., Allday, M. J. & Nair, V. (2006).** Interaction of MEQ protein and C-terminal-binding protein is critical for induction of lymphomas by Marek's disease virus. *Proc Natl Acad Sci U S A* **103**, 1687-1692.
- Brown, A. C., Smith, L. P., Kgosana, L., Baigent, S. J., Nair, V. & Allday, M. J. (2009).** Homodimerization of the Meq viral oncoprotein is necessary for induction of T-cell lymphoma by Marek's disease virus. *J Virol* **83**, 11142-11151.
- Brunovskis, P. & Velicer, L. F. (1995).** The Marek's disease virus (MDV) unique short region: alphaherpesvirus-homologous, fowlpox virus-homologous, and MDV-specific genes. *Virology* **206**, 324-338.
- Bryant, H. E., Matthews, D. A., Wadd, S., Scott, J. E., Kean, J., Graham, S., Russell, W. C. & Clements, J. B. (2000).** Interaction between herpes simplex virus type 1 IE63 protein and cellular protein p32. *J Virol* **74**, 11322-11328.
- Burgess, S. C., Young, J. R., Baaten, B. J., Hunt, L., Ross, L. N., Parcels, M. S., Kumar, P. M., Tregaskes, C. A., Lee, L. F. & Davison, T. F. (2004).** Marek's disease is a natural model for lymphomas overexpressing Hodgkin's disease antigen (CD30). *Proc Natl Acad Sci U S A* **101**, 13879-13884.
- Burley, S. K. & Roeder, R. G. (1996).** Biochemistry and structural biology of transcription factor IID (TFIID). *Annu Rev Biochem* **65**, 769-799.
- Burnside, J., Bernberg, E., Anderson, A., Lu, C., Meyers, B. C., Green, P. J., Jain, N., Isaacs, G. & Morgan, R. W. (2006).** Marek's disease virus encodes MicroRNAs that map to meq and the latency-associated transcript. *J Virol* **80**, 8778-8786.
- Burnside, J., Ouyang, M., Anderson, A., Bernberg, E., Lu, C., Meyers, B. C., Green, P. J., Markis, M., Isaacs, G., Huang, E. & Morgan, R. W. (2008).** Deep sequencing of chicken microRNAs. *BMC Genomics* **9**, 185.
- Burset, M., Seledtsov, I. A. & Solovyev, V. V. (2001).** SpliceDB: database of canonical and non-canonical mammalian splice sites. *Nucleic Acids Res* **29**, 255-259.
- Cai, X. & Cullen, B. R. (2006).** Transcriptional origin of Kaposi's sarcoma-associated herpesvirus microRNAs. *J Virol* **80**, 2234-2242.
- Cai, X., Schafer, A., Lu, S., Bilello, J. P., Desrosiers, R. C., Edwards, R., Raab-Traub, N. & Cullen, B. R. (2006).** Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed. *PLoS Pathog* **2**, e23.
- Calnek, B. W. (2001).** Pathogenesis of Marek's disease virus infection. *Curr Top Microbiol Immunol* **255**, 25-55.
- Calnek, B. W., Adldinger, H. K. & Kahn, D. E. (1970).** Feather follicle epithelium: a source of enveloped and infectious cell-free herpesvirus from Marek's disease. *Avian Dis* **14**, 219-233.
- Camp, H. S., Coussens, P. M. & Silva, R. F. (1991).** Cloning, sequencing, and functional analysis of a Marek's disease virus origin of DNA replication. *J Virol* **65**, 6320-6324.
- Cantello, J. L., Anderson, A. S. & Morgan, R. W. (1994).** Identification of latency-associated transcripts that map antisense to the ICP4 homolog gene of Marek's disease virus. *J Virol* **68**, 6280-6290.
- Cantello, J. L., Parcels, M. S., Anderson, A. S. & Morgan, R. W. (1997).** Marek's disease virus latency-associated transcripts belong to a family of spliced RNAs that are antisense to the ICP4 homolog gene. *J Virol* **71**, 1353-1361.
- Carpenter, J. A., Keegan, L. P., Wilfert, L., O'Connell, M. A. & Jiggins, F. M. (2009).** Evidence for ADAR-induced hypermutation of the Drosophila sigma virus (Rhabdoviridae). *BMC Genet* **10**, 75.
- Carthew, R. W. & Sontheimer, E. J. (2009).** Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-655.
- Ceppi, M., Pereira, P. M., Dunand-Sauthier, I., Barras, E., Reith, W., Santos, M. A. & Pierre, P. (2009).** MicroRNA-155 modulates the interleukin-1 signaling pathway in activated human monocyte-derived dendritic cells. *Proc Natl Acad Sci U S A* **106**, 2735-2740.

- Chang, K. S., Lee, S. I., Ohashi, K., Ibrahim, A. & Onuma, M. (2002a).** The detection of the meq gene in chicken infected with Marek's disease virus serotype 1. *J Vet Med Sci* **64**, 413-417.
- Chang, K. S., Ohashi, K. & Onuma, M. (2002b).** Diversity (polymorphism) of the meq gene in the attenuated Marek's disease virus (MDV) serotype 1 and MDV-transformed cell lines. *J Vet Med Sci* **64**, 1097-1101.
- Chang, K. S., Ohashi, K. & Onuma, M. (2002c).** Suppression of transcription activity of the MEQ protein of oncogenic Marek's disease virus serotype 1 (MDV1) by L-MEQ of non-oncogenic MDV1. *J Vet Med Sci* **64**, 1091-1095.
- Chbab, N., Chabanne-Vautherot, D., Francineau, A., Osterrieder, N., Denesvre, C. & Vautherot, J. F. (2009).** The Marek's disease virus (MDV) protein encoded by the UL17 ortholog is essential for virus growth. *Vet Res* **40**, 28.
- Chbab, N., Egerer, A., Veiga, I., Jarosinski, K. W. & Osterrieder, N. (2010).** Viral control of vTR expression is critical for efficient formation and dissemination of lymphoma induced by Marek's disease virus (MDV). *Vet Res* **41**, 56.
- Cheloufi, S., Dos Santos, C. O., Chong, M. M. & Hannon, G. J. (2010).** A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**, 584-589.
- Chen, R., Ding, J. & Wang, B. (2009).** The construction and characterization of the bi-directional promoter between pp38 gene and 1.8-kb mRNA transcripts of Marek's disease viruses. *Virol J* **6**, 212.
- Chen, S. H., Habib, G., Yang, C. Y., Gu, Z. W., Lee, B. R., Weng, S. A., Silberman, S. R., Cai, S. J., Deslypere, J. P., Rosseneu, M. & et al. (1987).** Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science* **238**, 363-366.
- Chendrimada, T. P., Finn, K. J., Ji, X., Baillat, D., Gregory, R. I., Liebhaber, S. A., Pasquinelli, A. E. & Shiekhattar, R. (2007).** MicroRNA silencing through RISC recruitment of eIF6. *Nature* **447**, 823-828.
- Chester, A., Weinreb, V., Carter, C. W., Jr. & Navaratnam, N. (2004).** Optimization of apolipoprotein B mRNA editing by APOBEC1 apoenzyme and the role of its auxiliary factor, ACF. *RNA* **10**, 1399-1411.
- Chi, J. H., Harley, C. A., Mukhopadhyay, A. & Wilson, D. W. (2005).** The cytoplasmic tail of herpes simplex virus envelope glycoprotein D binds to the tegument protein VP22 and to capsids. *J Gen Virol* **86**, 253-261.
- Cho, D. S., Yang, W., Lee, J. T., Shiekhattar, R., Murray, J. M. & Nishikura, K. (2003).** Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J Biol Chem* **278**, 17093-17102.
- Chou, J., Provot, S. & Werb, Z. (2010).** GATA3 in development and cancer differentiation: cells GATA have it! *J Cell Physiol* **222**, 42-49.
- Chouljenko, V. N., Iyer, A. V., Chowdhury, S., Kim, J. & Kousoulas, K. G. (2010).** The Herpes Simplex Virus Type-1 (HSV-1) UL20 Protein and the Amino Terminus of Glycoprotein K (gK) Physically Interact with Glycoprotein B (gB). *J Virol*.
- Choy, E. Y., Siu, K. L., Kok, K. H., Lung, R. W., Tsang, C. M., To, K. F., Kwong, D. L., Tsao, S. W. & Jin, D. Y. (2008).** An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival. *J Exp Med* **205**, 2551-2560.
- Churchill, A. E. & Biggs, P. M. (1967).** Agent of Marek's disease in tissue culture. *Nature* **215**, 528-530.
- Churchill, A. E., Chubb, R. C. & Baxendale, W. (1969).** The attenuation, with loss of oncogenicity, of the herpes-type virus of Marek's disease (strain HPRS-16) on passage in cell culture. *J Gen Virol* **4**, 557-564.
- Cifuentes, D., Xue, H., Taylor, D. W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G. J., Lawson, N. D., Wolfe, S. A. & Giraldez, A. J. (2010).** A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**, 1694-1698.
- Conticello, S. G. (2008).** The AID/APOBEC family of nucleic acid mutators. *Genome Biol* **9**, 229.
- Cooper, G. M. (2000).** *The cell*.

- Corcoran, D. L., Pandit, K. V., Gordon, B., Bhattacharjee, A., Kaminski, N. & Benos, P. V. (2009).** Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PLoS One* **4**, e5279.
- Cosman, D., Mullberg, J., Sutherland, C. L., Chin, W., Armitage, R., Fanslow, W., Kubin, M. & Chalupny, N. J. (2001).** ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. *Immunity* **14**, 123-133.
- Cui, X., Lee, L. F., Reed, W. M., Kung, H. J. & Reddy, S. M. (2004).** Marek's disease virus-encoded vIL-8 gene is involved in early cytolytic infection but dispensable for establishment of latency. *J Virol* **78**, 4753-4760.
- Cui, Z. Z., Yan, D. & Lee, L. F. (1990).** Marek's disease virus gene clones encoding virus-specific phosphorylated polypeptides and serological characterization of fusion proteins. *Virus Genes* **3**, 309-322.
- Darzacq, X., Jady, B. E., Verheggen, C., Kiss, A. M., Bertrand, E. & Kiss, T. (2002).** Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J* **21**, 2746-2756.
- Davie, J. R., He, S., Li, L., Sekhavat, A., Espino, P., Drohic, B., Dunn, K. L., Sun, J. M., Chen, H. Y., Yu, J., Pritchard, S. & Wang, X. (2008).** Nuclear organization and chromatin dynamics--Sp1, Sp3 and histone deacetylases. *Adv Enzyme Regul* **48**, 189-208.
- Davis, B. N. & Hata, A. (2009).** Regulation of MicroRNA Biogenesis: A miRiad of mechanisms. *Cell Commun Signal* **7**, 18.
- Davis, B. N., Hilyard, A. C., Lagna, G. & Hata, A. (2008).** SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **454**, 56-61.
- Davydov, I. V., Krammer, P. H. & Li-Weber, M. (1995).** Nuclear factor-IL6 activates the human IL-4 promoter in T cells. *J Immunol* **155**, 5273-5279.
- Dawson, T. R., Sansam, C. L. & Emeson, R. B. (2004).** Structure and sequence determinants required for the RNA editing of ADAR2 substrates. *J Biol Chem* **279**, 4941-4951.
- Debba-Pavard, M., Ait-Lounis, A., Soubieux, D., Rasschaert, D. & Dambrine, G. (2008a).** Vaccination against Marek's disease reduces telomerase activity and viral gene transcription in peripheral blood leukocytes from challenged chickens. *Vaccine* **26**, 4904-4912.
- Debba-Pavard, M., Le Galludec, H., Dambrine, G. & Rasschaert, D. (2008b).** Variations in the H/ACA box sequence of viral telomerase RNA of isolates of CVI988 Rispens vaccine. *Arch Virol* **153**, 1563-1568.
- Delecluse, H. J. & Hammerschmidt, W. (1993).** Status of Marek's disease virus in established lymphoma cell lines: herpesvirus integration is common. *J Virol* **67**, 82-92.
- Diederichs, S. & Haber, D. A. (2007).** Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **131**, 1097-1108.
- Dietrich, R. C., Fuller, J. D. & Padgett, R. A. (2005).** A mutational analysis of U12-dependent splice site dinucleotides. *RNA* **11**, 1430-1440.
- Ding, J., Cui, Z., Jiang, S. & Li, Y. (2008).** Study on the structure of heteropolymer pp38/pp24 and its enhancement on the bi-directional promoter upstream of pp38 gene in Marek's disease virus. *Sci China C Life Sci* **51**, 821-826.
- Ding, J., Cui, Z. & Lee, L. F. (2007).** Marek's disease virus unique genes pp38 and pp24 are essential for transactivating the bi-directional promoters for the 1.8 kb mRNA transcripts. *Virus Genes* **35**, 643-650.
- Ding, J., Cui, Z., Lee, L. F., Cui, X. & Reddy, S. M. (2006).** The role of pp38 in regulation of Marek's disease virus bi-directional promoter between pp38 and 1.8-kb mRNA. *Virus Genes* **32**, 193-201.
- Djeraba-AitLounis, A., Soubieux, D., Klapper, W. & Rasschaert, D. (2004).** Induction of telomerase activity in avian lymphoblastoid cell line transformed by Marek's disease virus, MDCC-MSB1. *Vet Pathol* **41**, 405-407.
- Dorange, F., Tischer, B. K., Vautherot, J. F. & Osterrieder, N. (2002).** Characterization of Marek's disease virus serotype 1 (MDV-1) deletion mutants that lack UL46 to UL49 genes: MDV-1 UL49, encoding VP22, is indispensable for virus growth. *J Virol* **76**, 1959-1970.

- Edwards-Gilbert, G., Veraldi, K. L. & Milcarek, C. (1997).** Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res* **25**, 2547-2561.
- Edwards, R. H., Marquitz, A. R. & Raab-Traub, N. (2008).** Epstein-Barr virus BART microRNAs are produced from a large intron prior to splicing. *J Virol* **82**, 9094-9106.
- Elliott, G., Mouzakis, G. & O'Hare, P. (1995).** VP16 interacts via its activation domain with VP22, a tegument protein of herpes simplex virus, and is relocated to a novel macromolecular assembly in coexpressing cells. *J Virol* **69**, 7932-7941.
- Elliott, G. & O'Hare, P. (1998).** Herpes simplex virus type 1 tegument protein VP22 induces the stabilization and hyperacetylation of microtubules. *J Virol* **72**, 6448-6455.
- Endoh, D., Ikegawa, S., Kon, Y., Hayashi, M. & Sato, F. (1995).** Expression of the endogenous Marek's disease virus ICP4 homolog (MDV ICP4) gene is enhanced in latently infected cells by transient transfection with the recombinant MDV ICP4 gene. *Jpn J Vet Res* **43**, 109-124.
- Eulalio, A., Behm-Ansmant, I. & Izaurralde, E. (2007).** P bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* **8**, 9-22.
- Eulalio, A., Huntzinger, E. & Izaurralde, E. (2008).** Getting to the root of miRNA-mediated gene silencing. *Cell* **132**, 9-14.
- Eulalio, A., Triteschler, F. & Izaurralde, E. (2009).** The GW182 protein family in animal cells: new insights into domains required for miRNA-mediated gene silencing. *RNA* **15**, 1433-1442.
- Fabian, M. R., Sonenberg, N. & Filipowicz, W. (2010).** Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* **79**, 351-379.
- Fabricant, J., Ianconescu, M. & Calnek, B. W. (1977).** Comparative effects of host and viral factors on early pathogenesis of Marek's disease. *Infect Immun* **16**, 136-144.
- Fang, Z. & Rajewsky, N. (2011).** The Impact of miRNA Target Sites in Coding Sequences and in 3'UTRs. *PLoS One* **6**, e18067.
- Felder, M. P., Laugier, D., Yatsula, B., Dezelee, P., Calothy, G. & Marx, M. (1994).** Functional and biological properties of an avian variant long terminal repeat containing multiple A to G conversions in the U3 sequence. *J Virol* **68**, 4759-4767.
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998).** Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811.
- Flint, J. & Shenk, T. (1997).** Viral transactivating proteins. *Annu Rev Genet* **31**, 177-212.
- Fox-Walsh, K. L., Dou, Y., Lam, B. J., Hung, S. P., Baldi, P. F. & Hertel, K. J. (2005).** The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proc Natl Acad Sci U S A* **102**, 16176-16181.
- Fragnet, L., Blasco, M. A., Klapper, W. & Rasschaert, D. (2003).** The RNA subunit of telomerase is encoded by Marek's disease virus. *J Virol* **77**, 5985-5996.
- Fragnet, L., Kut, E. & Rasschaert, D. (2005).** Comparative functional study of the viral telomerase RNA based on natural mutations. *J Biol Chem* **280**, 23502-23515.
- Frank, F., Sonenberg, N. & Nagar, B. (2010).** Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**, 818-822.
- Frilander, M. J. & Steitz, J. A. (1999).** Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. *Genes Dev* **13**, 851-863.
- Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T., Akimoto, C., Yamamoto, Y., Katagiri, T., Foulds, C., Takezawa, S., Kitagawa, H., Takeyama, K., O'Malley, B. W. & Kato, S. (2007).** DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. *Nat Cell Biol* **9**, 604-611.
- Gandy, S. Z., Linnstaedt, S. D., Muralidhar, S., Cashman, K. A., Rosenthal, L. J. & Casey, J. L. (2007).** RNA editing of the human herpesvirus 8 kaposin transcript eliminates its transforming activity and is induced during lytic replication. *J Virol* **81**, 13544-13551.
- Gimeno, I. M., Witter, R. L., Hunt, H. D., Reddy, S. M., Lee, L. F. & Silva, R. F. (2005).** The pp38 gene of Marek's disease virus (MDV) is necessary for cytolytic infection of B cells and maintenance of the transformed state but not for cytolytic infection of the feather follicle epithelium and horizontal spread of MDV. *J Virol* **79**, 4545-4549.

- Gonda, T. J. & Bishop, J. M. (1983).** Structure and transcription of the cellular homolog (c-myb) of the avian myeloblastosis virus transforming gene (v-myb). *J Virol* **46**, 212-220.
- Gott, J. M., Visomirski, L. M. & Hunter, J. L. (1993).** Substitutional and insertional RNA editing of the cytochrome c oxidase subunit 1 mRNA of *Physarum polycephalum*. *J Biol Chem* **268**, 25483-25486.
- Gottwein, E., Mukherjee, N., Sachse, C., Frenzel, C., Majoros, W. H., Chi, J. T., Braich, R., Manoharan, M., Soutschek, J., Ohler, U. & Cullen, B. R. (2007).** A viral microRNA functions as an orthologue of cellular miR-155. *Nature* **450**, 1096-1099.
- Graveley, B. R. (2001).** Alternative splicing: increasing diversity in the proteomic world. *Trends Genet* **17**, 100-107.
- Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N. & Shiekhattar, R. (2004).** The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-240.
- Grey, F., Meyers, H., White, E. A., Spector, D. H. & Nelson, J. (2007).** A human cytomegalovirus-encoded microRNA regulates expression of multiple viral genes involved in replication. *PLoS Pathog* **3**, e163.
- Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engle, P., Lim, L. P. & Bartel, D. P. (2007).** MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* **27**, 91-105.
- Gu, S., Jin, L., Zhang, F., Sarnow, P. & Kay, M. A. (2009).** Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nat Struct Mol Biol* **16**, 144-150.
- Guillon-Munos, A., Dambrine, G., Richerieux, N., Coupeau, D., Muylkens, B. & Rasschaert, D. (2010).** The chicken miR-150 targets the avian orthologue of the functional zebrafish MYB 3'UTR target site. *BMC Mol Biol* **11**, 67.
- Haase, A. D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A. & Filipowicz, W. (2005).** TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep* **6**, 961-967.
- Hagmann, M., Georgiev, O. & Schaffner, W. (1997).** The VP16 paradox: herpes simplex virus VP16 contains a long-range activation domain but within the natural multiprotein complex activates only from promoter-proximal positions. *J Virol* **71**, 5952-5962.
- Haj Khelil, A., Deguillien, M., Moriniere, M., Ben Chibani, J. & Baklouti, F. (2008).** Cryptic splicing sites are differentially utilized in vivo. *FEBS J* **275**, 1150-1162.
- Hajjar, A. M. & Linial, M. L. (1995).** Modification of retroviral RNA by double-stranded RNA adenosine deaminase. *J Virol* **69**, 5878-5882.
- Hall, S. L. & Padgett, R. A. (1994).** Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J Mol Biol* **239**, 357-365.
- Hansen, A., Henderson, S., Lagos, D., Nikitenko, L., Coulter, E., Roberts, S., Gratrix, F., Plaisance, K., Renne, R., Bower, M., Kellam, P. & Boshoff, C. (2010).** KSHV-encoded miRNAs target MAF to induce endothelial cell reprogramming. *Genes Dev* **24**, 195-205.
- Harris, R. S., Petersen-Mahrt, S. K. & Neuberger, M. S. (2002).** RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* **10**, 1247-1253.
- Hastings, M. L. & Krainer, A. R. (2001).** Functions of SR proteins in the U12-dependent AT-AC pre-mRNA splicing pathway. *RNA* **7**, 471-482.
- Hayashi, M., Kawamura, T., Akaike, H., Arai, S. & Okui, T. (1999).** Antisense oligonucleotide complementary to the BamHI-H gene family of Marek's disease virus induced growth arrest of MDCC-MSB1 cells in the S-phase. *J Vet Med Sci* **61**, 389-394.
- Heale, B. S., Keegan, L. P., McGurk, L., Michlewski, G., Brindle, J., Stanton, C. M., Caceres, J. F. & O'Connell, M. A. (2009).** Editing independent effects of ADARs on the miRNA/siRNA pathways. *EMBO J* **28**, 3145-3156.
- Heidari, M., Zhang, H. M. & Sharif, S. (2008).** Marek's disease virus induces Th-2 activity during cytolytic infection. *Viral Immunol* **21**, 203-214.
- Hengartner, C. J., Thompson, C. M., Zhang, J., Chao, D. M., Liao, S. M., Koleske, A. J., Okamura, S. & Young, R. A. (1995).** Association of an activator with an RNA polymerase II holoenzyme. *Genes Dev* **9**, 897-910.

- Henke, J. I., Goergen, D., Zheng, J., Song, Y., Schuttler, C. G., Fehr, C., Junemann, C. & Niepmann, M. (2008). microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* **27**, 3300-3310.
- Hernandez-Munain, C. & Krangel, M. S. (2002). Distinct roles for c-Myb and core binding factor/polyoma enhancer-binding protein 2 in the assembly and function of a multiprotein complex on the TCR delta enhancer in vivo. *J Immunol* **169**, 4362-4369.
- Hertel, K. J. (2008). Combinatorial control of exon recognition. *J Biol Chem* **283**, 1211-1215.
- Hicks, M. J., Yang, C. R., Kotlajich, M. V. & Hertel, K. J. (2006). Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol* **4**, e147.
- Higgs, M. R., Preston, V. G. & Stow, N. D. (2008). The UL15 protein of herpes simplex virus type 1 is necessary for the localization of the UL28 and UL33 proteins to viral DNA replication centres. *J Gen Virol* **89**, 1709-1715.
- Higuchi, M., Maas, S., Single, F. N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R. & Seeburg, P. H. (2000). Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406**, 78-81.
- Holmes, R. K., Malim, M. H. & Bishop, K. N. (2007). APOBEC-mediated viral restriction: not simply editing? *Trends Biochem Sci* **32**, 118-128.
- Hong, Y. & Coussens, P. M. (1994). Identification of an immediate-early gene in the Marek's disease virus long internal repeat region which encodes a unique 14-kilodalton polypeptide. *J Virol* **68**, 3593-3603.
- Hong, Y., Frame, M. & Coussens, P. M. (1995). A 14-kDa immediate-early phosphoprotein is specifically expressed in cells infected with oncogenic Marek's disease virus strains and their attenuated derivatives. *Virology* **206**, 695-700.
- Horton, T. L. & Landweber, L. F. (2000). Evolution of four types of RNA editing in myxomycetes. *RNA* **6**, 1339-1346.
- Horton, T. L. & Landweber, L. F. (2002). Rewriting the information in DNA: RNA editing in kinetoplasts and myxomycetes. *Curr Opin Microbiol* **5**, 620-626.
- Hsiang, Y. H., Goldman, J. P. & Raulet, D. H. (1995). The role of c-Myb or a related factor in regulating the T cell receptor gamma gene enhancer. *J Immunol* **154**, 5195-5204.
- Humphreys, D. T., Westman, B. J., Martin, D. I. & Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci U S A* **102**, 16961-16966.
- Hundley, H. A. & Bass, B. L. (2010). ADAR editing in double-stranded UTRs and other noncoding RNA sequences. *Trends Biochem Sci* **35**, 377-383.
- Hussain, M., Taft, R. J. & Asgari, S. (2008). An insect virus-encoded microRNA regulates viral replication. *J Virol* **82**, 9164-9170.
- Iizasa, H., Wulff, B. E., Alla, N. R., Maragkakis, M., Megraw, M., Hatzigeorgiou, A., Iwakiri, D., Takada, K., Wiedmer, A., Showe, L., Lieberman, P. & Nishikura, K. (2010). Editing of EBV-encoded BART6 microRNAs controls their dicer targeting and consequently affects viral latency. *J Biol Chem*.
- Islam, A. F., Walkden-Brown, S. W., Burgess, S. K. & Groves, P. J. (2001). Marek's disease in broiler chickens: effect of route of infection and herpesvirus of turkey-vaccination status on detection of virus from blood or spleen by polymerase chain reaction, and on weights of birds, bursa and spleen. *Avian Pathol* **30**, 621-628.
- Izumiya, Y., Jang, H. K., Ono, M. & Mikami, T. (2001). A complete genomic DNA sequence of Marek's disease virus type 2, strain HPRS24. *Curr Top Microbiol Immunol* **255**, 191-221.
- Jackson, I. J. (1991). A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res* **19**, 3795-3798.
- Jady, B. E., Darzacq, X., Tucker, K. E., Matera, A. G., Bertrand, E. & Kiss, T. (2003). Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO J* **22**, 1878-1888.
- Jarosinski, K. W., Margulis, N. G., Kamil, J. P., Spatz, S. J., Nair, V. K. & Osterrieder, N. (2007). Horizontal transmission of Marek's disease virus requires US2, the UL13 protein kinase, and gC. *J Virol* **81**, 10575-10587.

- Jarosinski, K. W., O'Connell, P. H. & Schat, K. A. (2003).** Impact of deletions within the Bam HI-L fragment of attenuated Marek's disease virus on vIL-8 expression and the newly identified transcript of open reading frame LORF4. *Virus Genes* **26**, 255-269.
- Jarosinski, K. W. & Osterrieder, N. (2010).** Further analysis of Marek's disease virus horizontal transmission confirms that U(L)44 (gC) and U(L)13 protein kinase activity are essential, while U(S)2 is nonessential. *J Virol* **84**, 7911-7916.
- Jarosinski, K. W., Osterrieder, N., Nair, V. K. & Schat, K. A. (2005).** Attenuation of Marek's disease virus by deletion of open reading frame RLORF4 but not RLORF5a. *J Virol* **79**, 11647-11659.
- Jarosinski, K. W. & Schat, K. A. (2006).** Multiple alternative splicing to exons II and III of viral interleukin-8 (vIL-8) in the Marek's disease virus genome: the importance of vIL-8 exon I. *Virus Genes* **34**, 9-22.
- Jepson, J. E. & Reenan, R. A. (2008).** RNA editing in regulating gene expression in the brain. *Biochim Biophys Acta* **1779**, 459-470.
- Johnston, M., Geoffroy, M. C., Sobala, A., Hay, R. & Hutvagner, G. (2010).** HSP90 protein stabilizes unloaded argonaute complexes and microscopically P-bodies in human cells. *Mol Biol Cell* **21**, 1462-1469.
- Jones, D., Lee, L., Liu, J. L., Kung, H. J. & Tillotson, J. K. (1992).** Marek disease virus encodes a basic-leucine zipper gene resembling the fos/jun oncogenes that is highly expressed in lymphoblastoid tumors. *Proc Natl Acad Sci U S A* **89**, 4042-4046.
- Jurak, I., Kramer, M. F., Mellor, J. C., van Lint, A. L., Roth, F. P., Knipe, D. M. & Coen, D. M. (2010).** Numerous conserved and divergent microRNAs expressed by herpes simplex viruses 1 and 2. *J Virol* **84**, 4659-4672.
- Jurica, M. S. & Moore, M. J. (2003).** Pre-mRNA splicing: awash in a sea of proteins. *Mol Cell* **12**, 5-14.
- Juven-Gershon, T. & Kadonaga, J. T. (2010).** Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol* **339**, 225-229.
- Kamil, J. P., Tischer, B. K., Trapp, S., Nair, V. K., Osterrieder, N. & Kung, H. J. (2005).** vLIP, a viral lipase homologue, is a virulence factor of Marek's disease virus. *J Virol* **79**, 6984-6996.
- Kang, J. W., Cho, S. H., Mo, I. P., Lee, D. W. & Kwon, H. J. (2007).** Prevalence and molecular characterization of meq in feather follicular epithelial cells of Korean broiler chickens. *Virus Genes* **35**, 339-345.
- Kastner, P. & Chan, S. (2008).** PU.1: a crucial and versatile player in hematopoiesis and leukemia. *Int J Biochem Cell Biol* **40**, 22-27.
- Kato, K., Izumiya, Y., Tohya, Y., Takahashi, E., Hirai, K. & Kawaguchi, Y. (2002).** Identification and characterization of Marek's disease virus serotype 1 (MDV1) ICP22 gene product: MDV1 ICP22 transactivates the MDV1 ICP27 promoter synergistically with MDV1 ICP4. *Vet Microbiol* **85**, 305-313.
- Katsumata, A., Iwata, A. & Ueda, S. (1998).** Cis-acting elements in the lytic origin of DNA replication of Marek's disease virus type 1. *J Gen Virol* **79** (Pt 12), 3015-3018.
- Kawaguchi, T., Nomura, K., Hirayama, Y. & Kitagawa, T. (1987).** Establishment and characterization of a chicken hepatocellular carcinoma cell line, LMH. *Cancer Res* **47**, 4460-4464.
- Kawahara, Y., Zinshteyn, B., Chendrimada, T. P., Shiekhhattar, R. & Nishikura, K. (2007a).** RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO Rep* **8**, 763-769.
- Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A. G. & Nishikura, K. (2007b).** Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **315**, 1137-1140.
- Kawamata, T. & Tomari, Y. (2010).** Making RISC. *Trends Biochem Sci* **35**, 368-376.
- Kawamura, M., Hayashi, M., Furuichi, T., Nonoyama, M., Isogai, E. & Namioka, S. (1991).** The inhibitory effects of oligonucleotides, complementary to Marek's disease virus mRNA transcribed from the BamHI-H region, on the proliferation of transformed lymphoblastoid cells, MDCC-MSB1. *J Gen Virol* **72** (Pt 5), 1105-1111.

- Kedde, M., Strasser, M. J., Boldajipour, B., Oude Vrielink, J. A., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P. M., van Duijse, J., Orom, U. A., Lund, A. H., Perrakis, A., Raz, E. & Agami, R. (2007). RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell* **131**, 1273-1286.
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U. & Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat Genet* **39**, 1278-1284.
- Kim, T., Hunt, H. D. & Cheng, H. H. (2010). Marek's disease viruses lacking either R-LORF10 or LORF4 have altered virulence in chickens. *Virus Genes* **40**, 410-420.
- Kishi, M., Bradley, G., Jessip, J., Tanaka, A. & Nonoyama, M. (1991). Inverted repeat regions of Marek's disease virus DNA possess a structure similar to that of the a sequence of herpes simplex virus DNA and contain host cell telomere sequences. *J Virol* **65**, 2791-2797.
- Klase, Z., Winograd, R., Davis, J., Carpio, L., Hildreth, R., Heydarian, M., Fu, S., McCaffrey, T., Meiri, E., Ayash-Rashkovsky, M., Gilad, S., Bentwich, Z. & Kashanchi, F. (2009). HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression. *Retrovirology* **6**, 18.
- Krol, J., Loedige, I. & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* **11**, 597-610.
- Kulkarni, S., Volchkova, V., Basler, C. F., Palese, P., Volchkov, V. E. & Shaw, M. L. (2009). Nipah virus edits its P gene at high frequency to express the V and W proteins. *J Virol* **83**, 3982-3987.
- Kumar, M. & Carmichael, G. G. (1997). Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci U S A* **94**, 3542-3547.
- Kunsch, C., Lang, R. K., Rosen, C. A. & Shannon, M. F. (1994). Synergistic transcriptional activation of the IL-8 gene by NF-kappa B p65 (RelA) and NF-IL-6. *J Immunol* **153**, 153-164.
- Kut, E. & Rasschaert, D. (2004). Assembly of Marek's disease virus (MDV) capsids using recombinant baculoviruses expressing MDV capsid proteins. *J Gen Virol* **85**, 769-774.
- Ladd, A. N. & Cooper, T. A. (2002). Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol* **3**, reviews0008.
- Lamberti, C. & Weller, S. K. (1998). The herpes simplex virus type 1 cleavage/packaging protein, UL32, is involved in efficient localization of capsids to replication compartments. *J Virol* **72**, 2463-2473.
- Landschulz, W. H., Johnson, P. F. & McKnight, S. L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759-1764.
- Laurent, S., Blondeau, C., Belghazi, M., Remy, S., Esnault, E., Rasschaert, P. & Rasschaert, D. (2007). Sequential autoprocessing of Marek's disease herpesvirus protease differs from that of other herpesviruses. *J Virol* **81**, 6117-6121.
- Le Rouzic, E. & Perbal, B. (1996). Retroviral insertional activation of the c-myc proto-oncogene in a Marek's disease T-lymphoma cell line. *J Virol* **70**, 7414-7423.
- Lee, L. F., Kreager, K. S., Arango, J., Paraguassu, A., Beckman, B., Zhang, H., Fadly, A., Lupiani, B. & Reddy, S. M. (2010). Comparative evaluation of vaccine efficacy of recombinant Marek's disease virus vaccine lacking Meq oncogene in commercial chickens. *Vaccine* **28**, 1294-1299.
- Lee, L. F., Lupiani, B., Silva, R. F., Kung, H. J. & Reddy, S. M. (2008). Recombinant Marek's disease virus (MDV) lacking the Meq oncogene confers protection against challenge with a very virulent plus strain of MDV. *Vaccine* **26**, 1887-1892.
- Lee, L. F., Wu, P., Sui, D., Ren, D., Kamil, J., Kung, H. J. & Witter, R. L. (2000). The complete unique long sequence and the overall genomic organization of the GA strain of Marek's disease virus. *Proc Natl Acad Sci U S A* **97**, 6091-6096.
- Lee, R. C., Feinbaum, R. L. & Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843-854.
- Lee, Y., Han, J., Yeom, K. H., Jin, H. & Kim, V. N. (2006). Drosha in primary microRNA processing. *Cold Spring Harb Symp Quant Biol* **71**, 51-57.
- Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H. & Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-4060.

- Lehmann, K. A. & Bass, B. L. (1999).** The importance of internal loops within RNA substrates of ADAR1. *J Mol Biol* **291**, 1-13.
- Lei, L., Ren, D. & Burton, Z. F. (1999).** The RAP74 subunit of human transcription factor IIF has similar roles in initiation and elongation. *Mol Cell Biol* **19**, 8372-8382.
- Lennartsson, A., Garwicz, D., Lindmark, A. & Gullberg, U. (2005).** The proximal promoter of the human cathepsin G gene conferring myeloid-specific expression includes C/EBP, c-myb and PU.1 binding sites. *Gene* **356**, 193-202.
- Levanon, E. Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z. Y., Shoshan, A., Pollock, S. R., Szybel, D., Olshansky, M., Rechavi, G. & Jantsch, M. F. (2004).** Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol* **22**, 1001-1005.
- Levine, A. & Durbin, R. (2001).** A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Res* **29**, 4006-4013.
- Levy, A. M., Burgess, S. C., Davidson, I., Underwood, G., Leitner, G. & Heller, E. D. (2003a).** Interferon-containing supernatants increase Marek's disease herpesvirus genomes and gene transcription levels, but not virion replication in vitro. *Viral Immunol* **16**, 501-509.
- Levy, A. M., Gilad, O., Xia, L., Izumiya, Y., Choi, J., Tsalenko, A., Yakhini, Z., Witter, R., Lee, L., Cardona, C. J. & Kung, H. J. (2005).** Marek's disease virus Meq transforms chicken cells via the v-Jun transcriptional cascade: a converging transforming pathway for avian oncoviruses. *Proc Natl Acad Sci U S A* **102**, 14831-14836.
- Levy, A. M., Izumiya, Y., Brunovskis, P., Xia, L., Parcells, M. S., Reddy, S. M., Lee, L., Chen, H. W. & Kung, H. J. (2003b).** Characterization of the chromosomal binding sites and dimerization partners of the viral oncoprotein Meq in Marek's disease virus-transformed T cells. *J Virol* **77**, 12841-12851.
- Lewis, B. P., Burge, C. B. & Bartel, D. P. (2005).** Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20.
- Li, D., O'Sullivan, G., Greenall, L., Smith, G., Jiang, C. & Ross, N. (1998).** Further characterization of the latency-associated transcription unit of Marek's disease virus. *Arch Virol* **143**, 295-311.
- Li, D. S., Pastorek, J., Zelnik, V., Smith, G. D. & Ross, L. J. (1994).** Identification of novel transcripts complementary to the Marek's disease virus homologue of the ICP4 gene of herpes simplex virus. *J Gen Virol* **75** (Pt 7), 1713-1722.
- Li, L. & Davie, J. R. (2010).** The role of Sp1 and Sp3 in normal and cancer cell biology. *Ann Anat* **192**, 275-283.
- Li, X., Jarosinski, K. W. & Schat, K. A. (2006).** Expression of Marek's disease virus phosphorylated polypeptide pp38 produces splice variants and enhances metabolic activity. *Vet Microbiol* **117**, 154-168.
- Liao, W., Hong, S. H., Chan, B. H., Rudolph, F. B., Clark, S. C. & Chan, L. (1999).** APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochem Biophys Res Commun* **260**, 398-404.
- Lin, C. F., Mount, S. M., Jarmolowski, A. & Makalowski, W. (2010).** Evolutionary dynamics of U12-type spliceosomal introns. *BMC Evol Biol* **10**, 47.
- Lin, J. & Cullen, B. R. (2007).** Analysis of the interaction of primate retroviruses with the human RNA interference machinery. *J Virol* **81**, 12218-12226.
- Liu, H. C., Soderblom, E. J. & Goshe, M. B. (2006).** A mass spectrometry-based proteomic approach to study Marek's Disease Virus gene expression. *J Virol Methods* **135**, 66-75.
- Liu, J. L. & Kung, H. J. (2000).** Marek's disease herpesvirus transforming protein MEQ: a c-Jun analogue with an alternative life style. *Virus Genes* **21**, 51-64.
- Liu, J. L., Lin, S. F., Xia, L., Brunovskis, P., Li, D., Davidson, I., Lee, L. F. & Kung, H. J. (1999a).** MEQ and V-IL8: cellular genes in disguise? *Acta Virol* **43**, 94-101.
- Liu, J. L., Ye, Y., Lee, L. F. & Kung, H. J. (1998).** Transforming potential of the herpesvirus oncoprotein MEQ: morphological transformation, serum-independent growth, and inhibition of apoptosis. *J Virol* **72**, 388-395.

- Liu, J. L., Ye, Y., Qian, Z., Qian, Y., Templeton, D. J., Lee, L. F. & Kung, H. J. (1999b).** Functional interactions between herpesvirus oncoprotein MEQ and cell cycle regulator CDK2. *J Virol* **73**, 4208-4219.
- Lo, A. K., To, K. F., Lo, K. W., Lung, R. W., Hui, J. W., Liao, G. & Hayward, S. D. (2007).** Modulation of LMP1 protein expression by EBV-encoded microRNAs. *Proc Natl Acad Sci U S A* **104**, 16164-16169.
- Long, J. C. & Caceres, J. F. (2009).** The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* **417**, 15-27.
- Long, M. C., Leong, V., Schaffer, P. A., Spencer, C. A. & Rice, S. A. (1999).** ICP22 and the UL13 protein kinase are both required for herpes simplex virus-induced modification of the large subunit of RNA polymerase II. *J Virol* **73**, 5593-5604.
- Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. (2004).** Nuclear export of microRNA precursors. *Science* **303**, 95-98.
- Lupiani, B., Lee, L. F. & Reddy, S. M. (2001).** Protein-coding content of the sequence of Marek's disease virus serotype 1. *Curr Top Microbiol Immunol* **255**, 159-190.
- Lutz, P. G., Houzel-Charavel, A., Moog-Lutz, C. & Cayre, Y. E. (2001).** Myeloblastin is an Myb target gene: mechanisms of regulation in myeloid leukemia cells growth-arrested by retinoic acid. *Blood* **97**, 2449-2456.
- Ma, J. B., Yuan, Y. R., Meister, G., Pei, Y., Tuschl, T. & Patel, D. J. (2005).** Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* **434**, 666-670.
- MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. & Doudna, J. A. (2008).** In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* **105**, 512-517.
- Maldonado, E., Ha, I., Cortes, P., Weis, L. & Reinberg, D. (1990).** Factors involved in specific transcription by mammalian RNA polymerase II: role of transcription factors IIA, IID, and IIB during formation of a transcription-competent complex. *Mol Cell Biol* **10**, 6335-6347.
- Marek, J. (1907).** Multiple Nervenentzündung (Polyneuritis) bei Hühnern. *Dtsch Tierärztl Wochenschr*, 417-421.
- Maris, C., Masse, J., Chester, A., Navaratnam, N. & Allain, F. H. (2005).** NMR structure of the apoB mRNA stem-loop and its interaction with the C to U editing APOBEC1 complementary factor. *RNA* **11**, 173-186.
- Maroney, P. A., Yu, Y., Fisher, J. & Nilsen, T. W. (2006).** Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat Struct Mol Biol* **13**, 1102-1107.
- Martinez-Contreras, R., Fiset, J. F., Nasim, F. U., Madden, R., Cordeau, M. & Chabot, B. (2006).** Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol* **4**, e21.
- McNab, A. R., Desai, P., Person, S., Roof, L. L., Thomsen, D. R., Newcomb, W. W., Brown, J. C. & Homa, F. L. (1998).** The product of the herpes simplex virus type 1 UL25 gene is required for encapsidation but not for cleavage of replicated viral DNA. *J Virol* **72**, 1060-1070.
- Mehta, A., Kinter, M. T., Sherman, N. E. & Driscoll, D. M. (2000).** Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol* **20**, 1846-1854.
- Meyer-Hoffert, U. (2009).** Neutrophil-derived serine proteases modulate innate immune responses. *Front Biosci* **14**, 3409-3418.
- Mikl, M. C., Watt, I. N., Lu, M., Reik, W., Davies, S. L., Neuberger, M. S. & Rada, C. (2005).** Mice deficient in APOBEC2 and APOBEC3. *Mol Cell Biol* **25**, 7270-7277.
- Modrek, B. & Lee, C. (2002).** A genomic view of alternative splicing. *Nat Genet* **30**, 13-19.
- Monteys, A. M., Spengler, R. M., Wan, J., Tecedor, L., Lennox, K. A., Xing, Y. & Davidson, B. L. (2010).** Structure and activity of putative intronic miRNA promoters. *RNA* **16**, 495-505.
- Moretti, F., Thermann, R. & Hentze, M. W. (2010).** Mechanism of translational regulation by miR-2 from sites in the 5' untranslated region or the open reading frame. *RNA* **16**, 2493-2502.
- Morgan, R., Anderson, A., Bernberg, E., Kamboj, S., Huang, E., Lagasse, G., Isaacs, G., Parcells, M., Meyers, B. C., Green, P. J. & Burnside, J. (2008).** Sequence conservation and differential expression of Marek's disease virus microRNAs. *J Virol* **82**, 12213-12220.

- Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I. & Proudfoot, N. J. (2008).** Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol* **15**, 902-909.
- Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. & Honjo, T. (2000).** Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553-563.
- Muramatsu, M., Sankaranand, V. S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O. & Honjo, T. (1999).** Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem* **274**, 18470-18476.
- Murphy, E., Vanicek, J., Robins, H., Shenk, T. & Levine, A. J. (2008).** Suppression of immediate-early viral gene expression by herpesvirus-coded microRNAs: implications for latency. *Proc Natl Acad Sci U S A* **105**, 5453-5458.
- Murre, C., McCaw, P. S. & Baltimore, D. (1989).** A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**, 777-783.
- Muylkens, B., Coupeau, D., Dambrine, G., Trapp, S. & Rasschaert, D. (2010).** Marek's disease virus microRNA designated Mdv1-pre-miR-M4 targets both cellular and viral genes. *Arch Virol* **155**, 1823-1837.
- Nagy, E. & Maquat, L. E. (1998).** A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci* **23**, 198-199.
- Nakajima, K., Ikuta, K., Naito, M., Ueda, S., Kato, S. & Hirai, K. (1987).** Analysis of Marek's disease virus serotype 1-specific phosphorylated polypeptides in virus-infected cells and Marek's disease lymphoblastoid cells. *J Gen Virol* **68** (Pt 5), 1379-1389.
- Narayanan, A., Nogueira, M. L., Ruyechan, W. T. & Kristie, T. M. (2005).** Combinatorial transcription of herpes simplex virus and varicella zoster virus immediate early genes is strictly determined by the cellular coactivator HCF-1. *J Biol Chem* **280**, 1369-1375.
- Nazerian, K. (1973).** Oncogenesis of Marek's disease. *Cancer Res* **33**, 1427-1430.
- Nazerian, K., Solomon, J. J., Witter, R. L. & Burmester, B. R. (1968).** Studies on the etiology of Marek's disease. II. Finding of a herpesvirus in cell culture. *Proc Soc Exp Biol Med* **127**, 177-182.
- Newcomb, W. W., Juhas, R. M., Thomsen, D. R., Homa, F. L., Burch, A. D., Weller, S. K. & Brown, J. C. (2001).** The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. *J Virol* **75**, 10923-10932.
- Newman, A. J. (1997).** The role of U5 snRNP in pre-mRNA splicing. *EMBO J* **16**, 5797-5800.
- Niikura, M., Kim, T., Hunt, H. D., Burnside, J., Morgan, R. W., Dodgson, J. B. & Cheng, H. H. (2007).** Marek's disease virus up-regulates major histocompatibility complex class II cell surface expression in infected cells. *Virology* **359**, 212-219.
- Niikura, M., Liu, H. C., Dodgson, J. B. & Cheng, H. H. (2004).** A comprehensive screen for chicken proteins that interact with proteins unique to virulent strains of Marek's disease virus. *Poult Sci* **83**, 1117-1123.
- Nishikura, K. (2010).** Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* **79**, 321-349.
- Nolde, M. J., Saka, N., Reinert, K. L. & Slack, F. J. (2007).** The *Caenorhabditis elegans* pumilio homolog, puf-9, is required for the 3'UTR-mediated repression of the let-7 microRNA target gene, hbl-1. *Dev Biol* **305**, 551-563.
- Nottrott, S., Simard, M. J. & Richter, J. D. (2006).** Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat Struct Mol Biol* **13**, 1108-1114.
- O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. (2005).** c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839-843.
- O'Hara, P. J., Nichol, S. T., Horodyski, F. M. & Holland, J. J. (1984).** Vesicular stomatitis virus defective interfering particles can contain extensive genomic sequence rearrangements and base substitutions. *Cell* **36**, 915-924.
- Oelgeschlager, M., Nuchprayoon, I., Luscher, B. & Friedman, A. D. (1996).** C/EBP, c-Myb, and PU.1 cooperate to regulate the neutrophil elastase promoter. *Mol Cell Biol* **16**, 4717-4725.

- Ogasawara, M., Suzutani, T., Yoshida, I. & Azuma, M. (2001). Role of the UL25 gene product in packaging DNA into the herpes simplex virus capsid: location of UL25 product in the capsid and demonstration that it binds DNA. *J Virol* **75**, 1427-1436.
- Ohashi, K., Zhou, W., O'Connell, P. H. & Schat, K. A. (1994). Characterization of a Marek's disease virus BamHI-L-specific cDNA clone obtained from a Marek's disease lymphoblastoid cell line. *J Virol* **68**, 1191-1195.
- Ohkuma, Y., Hashimoto, S., Wang, C. K., Horikoshi, M. & Roeder, R. G. (1995). Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE- α . *Mol Cell Biol* **15**, 4856-4866.
- Okada, T., Takagi, M., Murata, S., Onuma, M. & Ohashi, K. (2007). Identification and characterization of a novel spliced form of the meq transcript in lymphoblastoid cell lines derived from Marek's disease tumours. *J Gen Virol* **88**, 2111-2120.
- Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **130**, 89-100.
- Omar, A. R. & Schat, K. A. (1996). Syngeneic Marek's disease virus (MDV)-specific cell-mediated immune responses against immediate early, late, and unique MDV proteins. *Virology* **222**, 87-99.
- Omoto, S., Ito, M., Tsutsumi, Y., Ichikawa, Y., Okuyama, H., Brisibe, E. A., Saksena, N. K. & Fujii, Y. R. (2004). HIV-1 nef suppression by virally encoded microRNA. *Retrovirology* **1**, 44.
- Ono, M., Kawaguchi, Y., Maeda, K., Kamiya, N., Tohya, Y., Kai, C., Niikura, M. & Mikami, T. (1994). Nucleotide sequence analysis of Marek's disease virus (MDV) serotype 2 homolog of MDV serotype 1 pp38, an antigen associated with transformed cells. *Virology* **201**, 142-146.
- Orom, U. A., Nielsen, F. C. & Lund, A. H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell* **30**, 460-471.
- Osenberg, S., Paz Yaacov, N., Safran, M., Moshkovitz, S., Shtrichman, R., Sherf, O., Jacob-Hirsch, J., Keshet, G., Amariglio, N., Itskovitz-Eldor, J. & Rechavi, G. (2010). Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing. *PLoS One* **5**, e11173.
- Osterrieder, N., Kamil, J. P., Schumacher, D., Tischer, B. K. & Trapp, S. (2006). Marek's disease virus: from miasma to model. *Nat Rev Microbiol* **4**, 283-294.
- Pabo, C. O. & Sauer, R. T. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem* **61**, 1053-1095.
- Parcells, M. S., Anderson, A. S. & Morgan, T. W. (1995). Retention of oncogenicity by a Marek's disease virus mutant lacking six unique short region genes. *J Virol* **69**, 7888-7898.
- Parcells, M. S., Arumugaswami, V., Prigge, J. T., Pandya, K. & Dienglewicz, R. L. (2003). Marek's disease virus reactivation from latency: changes in gene expression at the origin of replication. *Poult Sci* **82**, 893-898.
- Parcells, M. S., Lin, S. F., Dienglewicz, R. L., Majerciak, V., Robinson, D. R., Chen, H. C., Wu, Z., Dubyak, G. R., Brunovskis, P., Hunt, H. D., Lee, L. F. & Kung, H. J. (2001). Marek's disease virus (MDV) encodes an interleukin-8 homolog (vIL-8): characterization of the vIL-8 protein and a vIL-8 deletion mutant MDV. *J Virol* **75**, 5159-5173.
- Parker, J. S., Roe, S. M. & Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* **434**, 663-666.
- Parker, M. A. & Schierman, L. W. (1983). Suppression of humoral immunity in chickens prevents transient paralysis caused by a herpesvirus. *J Immunol* **130**, 2000-2001.
- Paroo, Z., Ye, X., Chen, S. & Liu, Q. (2009). Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. *Cell* **139**, 112-122.
- Peng, F., Bradley, G., Tanaka, A., Lancz, G. & Nonoyama, M. (1992). Isolation and characterization of cDNAs from BamHI-H gene family RNAs associated with the tumorigenicity of Marek's disease virus. *J Virol* **66**, 7389-7396.
- Peng, Q. & Shirazi, Y. (1996a). Characterization of the protein product encoded by a splicing variant of the Marek's disease virus Eco-Q gene (Meq). *Virology* **226**, 77-82.

- Peng, Q. & Shirazi, Y. (1996b).** Isolation and characterization of Marek's disease virus (MDV) cDNAs from a MDV-transformed lymphoblastoid cell line: identification of an open reading frame antisense to the MDV Eco-Q protein (Meq). *Virology* **221**, 368-374.
- Peng, Q., Zeng, M., Bhuiyan, Z. A., Ubukata, E., Tanaka, A., Nonoyama, M. & Shirazi, Y. (1995).** Isolation and characterization of Marek's disease virus (MDV) cDNAs mapping to the BamHI-I2, BamHI-Q2, and BamHI-L fragments of the MDV genome from lymphoblastoid cells transformed and persistently infected with MDV. *Virology* **213**, 590-599.
- Peters, L. & Meister, G. (2007).** Argonaute proteins: mediators of RNA silencing. *Mol Cell* **26**, 611-623.
- Petersen-Mahrt, S. K., Estmer, C., Ohrmalm, C., Matthews, D. A., Russell, W. C. & Akusjarvi, G. (1999).** The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *EMBO J* **18**, 1014-1024.
- Petersen, C. P., Bordeleau, M. E., Pelletier, J. & Sharp, P. A. (2006).** Short RNAs repress translation after initiation in mammalian cells. *Mol Cell* **21**, 533-542.
- Petit, V., Guetard, D., Renard, M., Keriell, A., Sitbon, M., Wain-Hobson, S. & Vartanian, J. P. (2009).** Murine APOBEC1 is a powerful mutator of retroviral and cellular RNA in vitro and in vivo. *J Mol Biol* **385**, 65-78.
- Pfeffer, S., Sewer, A., Lagos-Quintana, M., Sheridan, R., Sander, C., Grasser, F. A., van Dyk, L. F., Ho, C. K., Shuman, S., Chien, M., Russo, J. J., Ju, J., Randall, G., Lindenbach, B. D., Rice, C. M., Simon, V., Ho, D. D., Zavolan, M. & Tuschl, T. (2005).** Identification of microRNAs of the herpesvirus family. *Nat Methods* **2**, 269-276.
- Pfeffer, S., Zavolan, M., Grasser, F. A., Chien, M., Russo, J. J., Ju, J., John, B., Enright, A. J., Marks, D., Sander, C. & Tuschl, T. (2004).** Identification of virus-encoded microRNAs. *Science* **304**, 734-736.
- Phuphuakrat, A., Kraiwong, R., Boonarkart, C., Lauhakirti, D., Lee, T. H. & Auewarakul, P. (2008).** Double-stranded RNA adenosine deaminases enhance expression of human immunodeficiency virus type 1 proteins. *J Virol* **82**, 10864-10872.
- Piepenbrink, M. S., Li, X., O'Connell, P. H. & Schat, K. A. (2009).** Marek's disease virus phosphorylated polypeptide pp38 alters transcription rates of mitochondrial electron transport and oxidative phosphorylation genes. *Virus Genes* **39**, 102-112.
- Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E. & Filipowicz, W. (2005).** Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* **309**, 1573-1576.
- Place, R. F., Li, L. C., Pookot, D., Noonan, E. J. & Dahiya, R. (2008).** MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proc Natl Acad Sci U S A* **105**, 1608-1613.
- Plachetka, A., Chayka, O., Wilczek, C., Melnik, S., Bonifer, C. & Klempnauer, K. H. (2008).** C/EBPbeta induces chromatin opening at a cell-type-specific enhancer. *Mol Cell Biol* **28**, 2102-2112.
- Polson, A. G., Bass, B. L. & Casey, J. L. (1996).** RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature* **380**, 454-456.
- Pope, R. M., Leutz, A. & Ness, S. A. (1994).** C/EBP beta regulation of the tumor necrosis factor alpha gene. *J Clin Invest* **94**, 1449-1455.
- Porter, I. M. & Stow, N. D. (2004).** Replication, recombination and packaging of amplicon DNA in cells infected with the herpes simplex virus type 1 alkaline nuclease null mutant ambUL12. *J Gen Virol* **85**, 3501-3510.
- Poulsen, H., Nilsson, J., Damgaard, C. K., Egebjerg, J. & Kjems, J. (2001).** CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domain. *Mol Cell Biol* **21**, 7862-7871.
- Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J. & Scott, J. (1987).** A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell* **50**, 831-840.
- Pratt, W. D., Cantello, J., Morgan, R. W. & Schat, K. A. (1994).** Enhanced expression of the Marek's disease virus-specific phosphoproteins after stable transfection of MSB-1 cells with the Marek's disease virus homologue of ICP4. *Virology* **201**, 132-136.

- Preston, C. M., Frame, M. C. & Campbell, M. E. (1988).** A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. *Cell* **52**, 425-434.
- Prochnow, C., Bransteitter, R. & Chen, X. S. (2009).** APOBEC deaminases-mutases with defensive roles for immunity. *Sci China C Life Sci* **52**, 893-902.
- Proudfoot, N. J., Furger, A. & Dye, M. J. (2002).** Integrating mRNA processing with transcription. *Cell* **108**, 501-512.
- Qi, H. H., Ongusaha, P. P., Myllyharju, J., Cheng, D., Pakkanen, O., Shi, Y., Lee, S. W. & Peng, J. (2008).** Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature* **455**, 421-424.
- Qian, Z., Brunovskis, P., Rauscher, F., 3rd, Lee, L. & Kung, H. J. (1995).** Transactivation activity of Meq, a Marek's disease herpesvirus bZIP protein persistently expressed in latently infected transformed T cells. *J Virol* **69**, 4037-4044.
- Rahadiani, N., Takakuwa, T., Tresnasari, K., Morii, E. & Aozasa, K. (2008).** Latent membrane protein-1 of Epstein-Barr virus induces the expression of B-cell integration cluster, a precursor form of microRNA-155, in B lymphoma cell lines. *Biochem Biophys Res Commun* **377**, 579-583.
- Rajcani, J. & Vojvodova, A. (1998).** The role of herpes simplex virus glycoproteins in the virus replication cycle. *Acta Virol* **42**, 103-118.
- Rebagliati, M. R. & Melton, D. A. (1987).** Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity. *Cell* **48**, 599-605.
- Reddy, S. M., Lupiani, B., Gimeno, I. M., Silva, R. F., Lee, L. F. & Witter, R. L. (2002).** Rescue of a pathogenic Marek's disease virus with overlapping cosmid DNAs: use of a pp38 mutant to validate the technology for the study of gene function. *Proc Natl Acad Sci U S A* **99**, 7054-7059.
- Reese, T. A., Xia, J., Johnson, L. S., Zhou, X., Zhang, W. & Virgin, H. W. (2010).** Identification of novel microRNA-like molecules generated from herpesvirus and host tRNA transcripts. *J Virol* **84**, 10344-10353.
- Rehwinkel, J., Behm-Ansmant, I., Gatfield, D. & Izaurralde, E. (2005).** A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* **11**, 1640-1647.
- Ren, D., Lee, L. F. & Coussens, P. M. (1994).** Identification and characterization of Marek's disease virus genes homologous to ICP27 and glycoprotein K of herpes simplex virus-1. *Virology* **204**, 242-250.
- Rino, J. & Carmo-Fonseca, M. (2009).** The spliceosome: a self-organized macromolecular machine in the nucleus? *Trends Cell Biol* **19**, 375-384.
- Rispens, B. H., van Vloten, H., Mastenbroek, N., Maas, H. J. & Schat, K. A. (1972a).** Control of Marek's disease in the Netherlands. I. Isolation of an avirulent Marek's disease virus (strain CVI 988) and its use in laboratory vaccination trials. *Avian Dis* **16**, 108-125.
- Rispens, B. H., van Vloten, H., Mastenbroek, N., Maas, J. L. & Schat, K. A. (1972b).** Control of Marek's disease in the Netherlands. II. Field trials on vaccination with an avirulent strain (CVI 988) of Marek's disease virus. *Avian Dis* **16**, 126-138.
- Rodriguez, A., Vigorito, E., Clare, S., Warren, M. V., Couttet, P., Soond, D. R., van Dongen, S., Grocock, R. J., Das, P. P., Miska, E. A., Vetrie, D., Okkenhaug, K., Enright, A. J., Dougan, G., Turner, M. & Bradley, A. (2007).** Requirement of bic/microRNA-155 for normal immune function. *Science* **316**, 608-611.
- Rogozin, I. B., Basu, M. K., Jordan, I. K., Pavlov, Y. I. & Koonin, E. V. (2005).** APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. *Cell Cycle* **4**, 1281-1285.
- Roizman, B., Knipe, D. M. & Whitley, R. J. (2007).** Herpes Simplex Viruses. In *Fields Virology*, pp. 2501-2601. Edited by D. M. Knipe & P. M. Howley: Lippincott Williams & Wilkins.
- Rosati, M., Valentin, A., Patenaude, D. J. & Pavlakis, G. N. (2001).** CCAAT-enhancer-binding protein beta (C/EBP beta) activates CCR5 promoter: increased C/EBP beta and CCR5 in T lymphocytes from HIV-1-infected individuals. *J Immunol* **167**, 1654-1662.
- Roscigno, R. F. & Garcia-Blanco, M. A. (1995).** SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* **1**, 692-706.

- Ross, N., O'Sullivan, G., Rothwell, C., Smith, G., Burgess, S. C., Rennie, M., Lee, L. F. & Davison, T. F. (1997). Marek's disease virus EcoRI-Q gene (meq) and a small RNA antisense to ICP4 are abundantly expressed in CD4+ cells and cells carrying a novel lymphoid marker, AV37, in Marek's disease lymphomas. *J Gen Virol* **78** (Pt 9), 2191-2198.
- Rouha, H., Thurner, C. & Mandl, C. W. (2010). Functional microRNA generated from a cytoplasmic RNA virus. *Nucleic Acids Res.*
- Rueda, P., Garcia-Barreno, B. & Melero, J. A. (1994). Loss of conserved cysteine residues in the attachment (G) glycoprotein of two human respiratory syncytial virus escape mutants that contain multiple A-G substitutions (hypermutations). *Virology* **198**, 653-662.
- Rybak, A., Fuchs, H., Hadian, K., Smirnova, L., Wulczyn, E. A., Michel, G., Nitsch, R., Krappmann, D. & Wulczyn, F. G. (2009). The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. *Nat Cell Biol* **11**, 1411-1420.
- Sakamoto, S., Aoki, K., Higuchi, T., Todaka, H., Morisawa, K., Tamaki, N., Hatano, E., Fukushima, A., Taniguchi, T. & Agata, Y. (2009). The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. *Mol Cell Biol* **29**, 3754-3769.
- Sakurai, M., Yano, T., Kawabata, H., Ueda, H. & Suzuki, T. (2010). Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat Chem Biol* **6**, 733-740.
- Salmon, B., Cunningham, C., Davison, A. J., Harris, W. J. & Baines, J. D. (1998). The herpes simplex virus type 1 U(L)17 gene encodes virion tegument proteins that are required for cleavage and packaging of viral DNA. *J Virol* **72**, 3779-3788.
- Samols, M. A., Skalsky, R. L., Maldonado, A. M., Riva, A., Lopez, M. C., Baker, H. V. & Renne, R. (2007). Identification of cellular genes targeted by KSHV-encoded microRNAs. *PLoS Pathog* **3**, e65.
- Samuel, C. E. (2011). Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology* **411**, 180-193.
- Sanchez, A., Trappier, S. G., Mahy, B. W., Peters, C. J. & Nichol, S. T. (1996). The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc Natl Acad Sci U S A* **93**, 3602-3607.
- Sarnow, P., Jopling, C. L., Norman, K. L., Schutz, S. & Wehner, K. A. (2006). MicroRNAs: expression, avoidance and subversion by vertebrate viruses. *Nat Rev Microbiol* **4**, 651-659.
- Scadden, A. D. & O'Connell, M. A. (2005). Cleavage of dsRNAs hyper-edited by ADARs occurs at preferred editing sites. *Nucleic Acids Res* **33**, 5954-5964.
- Schat, K. A., Buckmaster, A. & Ross, L. J. (1989). Partial transcription map of Marek's disease herpesvirus in lytically infected cells and lymphoblastoid cell lines. *Int J Cancer* **44**, 101-109.
- Schat, K. A., Hooft van Iddekinge, B. J., Boerrigter, H., O'Connell, P. H. & Koch, G. (1998). Open reading frame L1 of Marek's disease herpesvirus is not essential for in vitro and in vivo virus replication and establishment of latency. *J Gen Virol* **79** (Pt 4), 841-849.
- Schat, K. A. & Nair, V. (2008a). Marek's disease. In *Disease of poultry*, pp. 452-514.
- Schat, K. A. & Nair, V. (2008b). Marek's Disease. In *Diseases of Poultry*, pp. 452-514. Edited by Y. M. Saif: Wiley-Blackwell.
- Schat, K. A. & Xing, Z. (2000). Specific and nonspecific immune responses to Marek's disease virus. *Dev Comp Immunol* **24**, 201-221.
- Schnall-Levin, M., Zhao, Y., Perrimon, N. & Berger, B. (2010). Conserved microRNA targeting in Drosophila is as widespread in coding regions as in 3'UTRs. *Proc Natl Acad Sci U S A* **107**, 15751-15756.
- Schneider, C., Will, C. L., Makarova, O. V., Makarov, E. M. & Luhrmann, R. (2002). Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Mol Cell Biol* **22**, 3219-3229.
- Schumacher, D., Tischer, B. K., Fuchs, W. & Osterrieder, N. (2000). Reconstitution of Marek's disease virus serotype 1 (MDV-1) from DNA cloned as a bacterial artificial chromosome and characterization of a glycoprotein B-negative MDV-1 mutant. *J Virol* **74**, 11088-11098.
- Schumacher, D., Tischer, B. K., Reddy, S. M. & Osterrieder, N. (2001). Glycoproteins E and I of Marek's disease virus serotype 1 are essential for virus growth in cultured cells. *J Virol* **75**, 11307-11318.

- Seo, G. J., Fink, L. H., O'Hara, B., Atwood, W. J. & Sullivan, C. S. (2008). Evolutionarily conserved function of a viral microRNA. *J Virol* **82**, 9823-9828.
- Seto, E., Moosmann, A., Gromminger, S., Walz, N., Grundhoff, A. & Hammerschmidt, W. (2010). Micro RNAs of Epstein-Barr virus promote cell cycle progression and prevent apoptosis of primary human B cells. *PLoS Pathog* **6**.
- Severini, A., Scraba, D. G. & Tyrrell, D. L. (1996). Branched structures in the intracellular DNA of herpes simplex virus type 1. *J Virol* **70**, 3169-3175.
- Shamblin, C. E., Greene, N., Arumugaswami, V., Dienglewicz, R. L. & Parcells, M. S. (2004). Comparative analysis of Marek's disease virus (MDV) glycoprotein-, lytic antigen pp38- and transformation antigen Meq-encoding genes: association of meq mutations with MDVs of high virulence. *Vet Microbiol* **102**, 147-167.
- Shao, L., Rapp, L. M. & Weller, S. K. (1993). Herpes simplex virus 1 alkaline nuclease is required for efficient egress of capsids from the nucleus. *Virology* **196**, 146-162.
- Shek, W. R., Calnek, B. W., Schat, K. A. & Chen, C. H. (1983). Characterization of Marek's disease virus-infected lymphocytes: discrimination between cytolytically and latently infected cells. *J Natl Cancer Inst* **70**, 485-491.
- Shen, W., Sa e Silva, M., Jaber, T., Vitvitskaia, O., Li, S., Henderson, G. & Jones, C. (2009). Two small RNAs encoded within the first 1.5 kilobases of the herpes simplex virus type 1 latency-associated transcript can inhibit productive infection and cooperate to inhibit apoptosis. *J Virol* **83**, 9131-9139.
- Shigekane, H., Kawaguchi, Y., Shirakata, M., Sakaguchi, M. & Hirai, K. (1999). The bi-directional transcriptional promoters for the latency-relating transcripts of the pp38/pp24 mRNAs and the 1.8 kb-mRNA in the long inverted repeats of Marek's disease virus serotype 1 DNA are regulated by common promoter-specific enhancers. *Arch Virol* **144**, 1893-1907.
- Shin, C., Nam, J. W., Farh, K. K., Chiang, H. R., Shkumatava, A. & Bartel, D. P. (2010). Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell* **38**, 789-802.
- Shkreli, M., Dambrine, G., Soubieux, D., Kut, E. & Rasschaert, D. (2007). Involvement of the oncoprotein c-Myc in viral telomerase RNA gene regulation during Marek's disease virus-induced lymphomagenesis. *J Virol* **81**, 4848-4857.
- Shukla, G. C. & Padgett, R. A. (2001). The intramolecular stem-loop structure of U6 snRNA can functionally replace the U6atac snRNA stem-loop. *RNA* **7**, 94-105.
- Silva, R. F. & Gimeno, I. (2007). Oncogenic Marek's disease viruses lacking the 132 base pair repeats can still be attenuated by serial in vitro cell culture passages. *Virus Genes* **34**, 87-90.
- Silva, R. F., Reddy, S. M. & Lupiani, B. (2004). Expansion of a unique region in the Marek's disease virus genome occurs concomitantly with attenuation but is not sufficient to cause attenuation. *J Virol* **78**, 733-740.
- Singh, J., Singh, C. P., Bhavani, A. & Nagaraju, J. (2010). Discovering microRNAs from Bombyx mori nucleopolyhedrosis virus. *Virology* **407**, 120-128.
- Skalsky, R. L., Samols, M. A., Plaisance, K. B., Boss, I. W., Riva, A., Lopez, M. C., Baker, H. V. & Renne, R. (2007). Kaposi's sarcoma-associated herpesvirus encodes an ortholog of miR-155. *J Virol* **81**, 12836-12845.
- Slezak-Prochazka, I., Durmus, S., Kroesen, B. J. & van den Berg, A. (2010). MicroRNAs, macrocontrol: regulation of miRNA processing. *RNA* **16**, 1087-1095.
- Smith, G. D., Zelnik, V. & Ross, L. J. (1995). Gene organization in herpesvirus of turkeys: identification of a novel open reading frame in the long unique region and a truncated homologue of pp38 in the internal repeat. *Virology* **207**, 205-216.
- Sodeik, B., Ebersold, M. W. & Helenius, A. (1997). Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. *J Cell Biol* **136**, 1007-1021.
- Spatz, S. J. (2010). Accumulation of attenuating mutations in varying proportions within a high passage very virulent plus strain of Gallid herpesvirus type 2. *Virus Res* **149**, 135-142.
- Spatz, S. J. & Silva, R. F. (2007a). Polymorphisms in the repeat long regions of oncogenic and attenuated pathotypes of Marek's disease virus 1. *Virus Genes* **35**, 41-53.
- Spatz, S. J. & Silva, R. F. (2007b). Sequence determination of variable regions within the genomes of gallid herpesvirus-2 pathotypes. *Arch Virol* **152**, 1665-1678.

- Spatz, S. J., Zhao, Y., Petherbridge, L., Smith, L. P., Baigent, S. J. & Nair, V. (2007).** Comparative sequence analysis of a highly oncogenic but horizontal spread-defective clone of Marek's disease virus. *Virus Genes* **35**, 753-766.
- Stern-Ginossar, N., Elefant, N., Zimmermann, A., Wolf, D. G., Saleh, N., Biton, M., Horwitz, E., Prokocimer, Z., Prichard, M., Hahn, G., Goldman-Wohl, D., Greenfield, C., Yagel, S., Hengel, H., Altuvia, Y., Margalit, H. & Mandelboim, O. (2007).** Host immune system gene targeting by a viral miRNA. *Science* **317**, 376-381.
- Stevens, S. W., Ryan, D. E., Ge, H. Y., Moore, R. E., Young, M. K., Lee, T. D. & Abelson, J. (2002).** Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol Cell* **9**, 31-44.
- Stik, G., Laurent, S., Coupeau, D., Coutaud, B., Dambrine, G., Rasschaert, D. & Muylkens, B. (2010).** A p53-dependent promoter associated with polymorphic tandem repeats controls the expression of a viral transcript encoding clustered microRNAs. *RNA*.
- Stuart, K. D., Schnaufer, A., Ernst, N. L. & Panigrahi, A. K. (2005).** Complex management: RNA editing in trypanosomes. *Trends Biochem Sci* **30**, 97-105.
- Suchodolski, P. F., Izumiya, Y., Lupiani, B., Ajithdoss, D. K., Gilad, O., Lee, L. F., Kung, H. J. & Reddy, S. M. (2009).** Homodimerization of Marek's disease virus-encoded Meq protein is not sufficient for transformation of lymphocytes in chickens. *J Virol* **83**, 859-869.
- Sullivan, C. S. (2008).** New roles for large and small viral RNAs in evading host defences. *Nat Rev Genet* **9**, 503-507.
- Sullivan, C. S., Grundhoff, A. T., Tevethia, S., Pipas, J. M. & Ganem, D. (2005).** SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* **435**, 682-686.
- Tabara, H., Grishok, A. & Mello, C. C. (1998).** RNAi in *C. elegans*: soaking in the genome sequence. *Science* **282**, 430-431.
- Tahiri-Alaoui, A., Matsuda, D., Xu, H., Panagiotis, P., Burman, L., Lambeth, L. S., Petherbridge, L., James, W., Mauro, V. & Nair, V. (2009a).** The 5' leader of the mRNA encoding the marek's disease virus serotype 1 pp14 protein contains an intronic internal ribosome entry site with allosteric properties. *J Virol* **83**, 12769-12778.
- Tahiri-Alaoui, A., Smith, L. P., Baigent, S., Kgosana, L., Petherbridge, L. J., Lambeth, L. S., James, W. & Nair, V. (2009b).** Identification of an intercistronic internal ribosome entry site in a Marek's disease virus immediate-early gene. *J Virol* **83**, 5846-5853.
- Tahirov, T. H., Sato, K., Ichikawa-Iwata, E., Sasaki, M., Inoue-Bungo, T., Shiina, M., Kimura, K., Takata, S., Fujikawa, A., Morii, H., Kumasaka, T., Yamamoto, M., Ishii, S. & Ogata, K. (2002).** Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter. *Cell* **108**, 57-70.
- Tan, S. & Richmond, T. J. (1998).** Eukaryotic transcription factors. *Curr Opin Struct Biol* **8**, 41-48.
- Tang, S., Bertke, A. S., Patel, A., Wang, K., Cohen, J. I. & Krause, P. R. (2008).** An acutely and latently expressed herpes simplex virus 2 viral microRNA inhibits expression of ICP34.5, a viral neurovirulence factor. *Proc Natl Acad Sci U S A* **105**, 10931-10936.
- Tang, S., Patel, A. & Krause, P. R. (2009).** Novel less-abundant viral microRNAs encoded by herpes simplex virus 2 latency-associated transcript and their roles in regulating ICP34.5 and ICP0 mRNAs. *J Virol* **83**, 1433-1442.
- Tange, T. O., Damgaard, C. K., Guth, S., Valcarcel, J. & Kjems, J. (2001).** The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO J* **20**, 5748-5758.
- Tarn, W. Y. & Steitz, J. A. (1995).** Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. *Proc Natl Acad Sci U S A* **92**, 2504-2508.
- Tatman, J. D., Preston, V. G., Nicholson, P., Elliott, R. M. & Rixon, F. J. (1994).** Assembly of herpes simplex virus type 1 capsids using a panel of recombinant baculoviruses. *J Gen Virol* **75** (Pt 5), 1101-1113.
- Taus, N. S., Salmon, B. & Baines, J. D. (1998).** The herpes simplex virus 1 UL 17 gene is required for localization of capsids and major and minor capsid proteins to intranuclear sites where viral DNA is cleaved and packaged. *Virology* **252**, 115-125.

- Tay, Y., Zhang, J., Thomson, A. M., Lim, B. & Rigoutsos, I. (2008).** MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* **455**, 1124-1128.
- Timmons, L. & Fire, A. (1998).** Specific interference by ingested dsRNA. *Nature* **395**, 854.
- Tischer, B. K., Schumacher, D., Messerle, M., Wagner, M. & Osterrieder, N. (2002).** The products of the UL10 (gM) and the UL49.5 genes of Marek's disease virus serotype 1 are essential for virus growth in cultured cells. *J Gen Virol* **83**, 997-1003.
- Toth, A. M., Li, Z., Cattaneo, R. & Samuel, C. E. (2009).** RNA-specific adenosine deaminase ADAR1 suppresses measles virus-induced apoptosis and activation of protein kinase PKR. *J Biol Chem* **284**, 29350-29356.
- Trapp, S., Parcells, M. S., Kamil, J. P., Schumacher, D., Tischer, B. K., Kumar, P. M., Nair, V. K. & Osterrieder, N. (2006).** A virus-encoded telomerase RNA promotes malignant T cell lymphomagenesis. *J Exp Med* **203**, 1307-1317.
- Tritschler, F., Huntzinger, E. & Izaurralde, E. (2010).** Role of GW182 proteins and PABPC1 in the miRNA pathway: a sense of déjà vu. *Nat Rev Mol Cell Biol* **11**, 379-384.
- Tulman, E. R., Afonso, C. L., Lu, Z., Zsak, L., Rock, D. L. & Kutish, G. F. (2000).** The genome of a very virulent Marek's disease virus. *J Virol* **74**, 7980-7988.
- Turunen, J. J., Will, C. L., Grote, M., Luhrmann, R. & Frilander, M. J. (2008).** The U11-48K protein contacts the 5' splice site of U12-type introns and the U11-59K protein. *Mol Cell Biol* **28**, 3548-3560.
- Umbach, J. L., Kramer, M. F., Jurak, I., Karnowski, H. W., Coen, D. M. & Cullen, B. R. (2008).** MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. *Nature* **454**, 780-783.
- Valente, L. & Nishikura, K. (2007).** RNA binding-independent dimerization of adenosine deaminases acting on RNA and dominant negative effects of nonfunctional subunits on dimer functions. *J Biol Chem* **282**, 16054-16061.
- van Iddekinge, B. J., Stenzler, L., Schat, K. A., Boerrigter, H. & Koch, G. (1999).** Genome analysis of Marek's disease virus strain CVI-988: effect of cell culture passage on the inverted repeat regions. *Avian Dis* **43**, 182-188.
- Vargova, K., Curik, N., Burda, P., Basova, P., Kulvait, V., Pospisil, V., Savvulidi, F., Kokavec, J., Necas, E., Berkova, A., Obrtlíkova, P., Karban, J., Mraz, M., Pospisilova, S., Mayer, J., Trneny, M., Zavadil, J. & Stopka, T. (2011).** MYB transcriptionally regulates the miR-155 host gene in chronic lymphocytic leukemia. *Blood* **117**, 3816-3825.
- Vasudevan, S., Tong, Y. & Steitz, J. A. (2007).** Switching from repression to activation: microRNAs can up-regulate translation. *Science* **318**, 1931-1934.
- Vasudevan, S., Tong, Y. & Steitz, J. A. (2008).** Cell-cycle control of microRNA-mediated translation regulation. *Cell Cycle* **7**, 1545-1549.
- Vigorito, E., Perks, K. L., Abreu-Goodger, C., Bunting, S., Xiang, Z., Kohlhaas, S., Das, P. P., Miska, E. A., Rodriguez, A., Bradley, A., Smith, K. G., Rada, C., Enright, A. J., Toellner, K. M., MacLennan, I. C. & Turner, M. (2007).** microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity* **27**, 847-859.
- Voinnet, O. (2009).** Origin, biogenesis, and activity of plant microRNAs. *Cell* **136**, 669-687.
- Volchkov, V. E., Becker, S., Volchkova, V. A., Ternovoj, V. A., Kotov, A. N., Netesov, S. V. & Klenk, H. D. (1995).** GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. *Virology* **214**, 421-430.
- Waidner, L. A., Morgan, R. W., Anderson, A. S., Bernberg, E. L., Kamboj, S., Garcia, M., Riblet, S. M., Ouyang, M., Isaacs, G. K., Markis, M., Meyers, B. C., Green, P. J. & Burnside, J. (2009).** MicroRNAs of Gallid and Meleagrid herpesviruses show generally conserved genomic locations and are virus-specific. *Virology* **388**, 128-136.
- Wang, B., Yanez, A. & Novina, C. D. (2008a).** MicroRNA-repressed mRNAs contain 40S but not 60S components. *Proc Natl Acad Sci U S A* **105**, 5343-5348.
- Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T. & Patel, D. J. (2008b).** Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921-926.

- Weinheimer, S. P., Boyd, B. A., Durham, S. K., Resnick, J. L. & O'Boyle, D. R., 2nd (1992). Deletion of the VP16 open reading frame of herpes simplex virus type 1. *J Virol* **66**, 258-269.
- White, C. A., Stow, N. D., Patel, A. H., Hughes, M. & Preston, V. G. (2003). Herpes simplex virus type 1 portal protein UL6 interacts with the putative terminase subunits UL15 and UL28. *J Virol* **77**, 6351-6358.
- Will, C. L. & Luhrmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13**, 290-301.
- Witter, R. L. (1987). New serotype 2 and attenuated serotype 1 Marek's disease vaccine viruses: comparative efficacy. *Avian Dis* **31**, 752-765.
- Witter, R. L., Li, D., Jones, D., Lee, L. F. & Kung, H. J. (1997). Retroviral insertional mutagenesis of a herpesvirus: a Marek's disease virus mutant attenuated for oncogenicity but not for immunosuppression or in vivo replication. *Avian Dis* **41**, 407-421.
- Witter, R. L., Nazerian, K., Purchase, H. G. & Burgoyne, G. H. (1970). Isolation from turkeys of a cell-associated herpesvirus antigenically related to Marek's disease virus. *Am J Vet Res* **31**, 525-538.
- Witter, R. L., Solomon, J. J., Champion, L. R. & Nazerian, K. (1971). Long-term studies of Marek's disease infection in individual chickens. *Avian Dis* **15**, 346-365.
- Wolf, J., Gerber, A. P. & Keller, W. (2002). tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J* **21**, 3841-3851.
- Wong, S. K. & Lazinski, D. W. (2002). Replicating hepatitis delta virus RNA is edited in the nucleus by the small form of ADAR1. *Proc Natl Acad Sci U S A* **99**, 15118-15123.
- Wu, P., Reed, W. M. & Lee, L. F. (2001). Glycoproteins H and L of Marek's disease virus form a hetero-oligomer essential for translocation and cell surface expression. *Arch Virol* **146**, 983-992.
- Wu, P., Reed, W. M., Yoshida, S., Sui, D. & Lee, L. F. (1999). Identification and characterization of glycoprotein H of MDV-1 GA strain. *Acta Virol* **43**, 152-158.
- Xia, T., O'Hara, A., Araujo, I., Barreto, J., Carvalho, E., Sapucaia, J. B., Ramos, J. C., Luz, E., Pedroso, C., Manrique, M., Toomey, N. L., Brites, C., Dittmer, D. P. & Harrington, W. J., Jr. (2008). EBV microRNAs in primary lymphomas and targeting of CXCL-11 by ebv-mir-BHRF1-3. *Cancer Res* **68**, 1436-1442.
- Xiang, J. & Wu, J. (2010). Feud or Friend? The Role of the miR-17-92 Cluster in Tumorigenesis. *Curr Genomics* **11**, 129-135.
- Xie, Q., Anderson, A. S. & Morgan, R. W. (1996). Marek's disease virus (MDV) ICP4, pp38, and meq genes are involved in the maintenance of transformation of MDCC-MSB1 MDV-transformed lymphoblastoid cells. *J Virol* **70**, 1125-1131.
- Xing, L. & Kieff, E. (2007). Epstein-Barr virus BHRF1 micro- and stable RNAs during latency III and after induction of replication. *J Virol* **81**, 9967-9975.
- Xing, Z. & Schat, K. A. (2000). Inhibitory effects of nitric oxide and gamma interferon on in vitro and in vivo replication of Marek's disease virus. *J Virol* **74**, 3605-3612.
- Xu, S., Xue, C., Li, J., Bi, Y. & Cao, Y. (2011). Marek's disease virus type 1 microRNA miR-M3 suppresses cisplatin-induced apoptosis by targeting Smad2 of the transforming growth factor beta signal pathway. *J Virol* **85**, 276-285.
- Yamagata, K., Fujiyama, S., Ito, S., Ueda, T., Murata, T., Naitou, M., Takeyama, K., Minami, Y., O'Malley, B. W. & Kato, S. (2009). Maturation of microRNA is hormonally regulated by a nuclear receptor. *Mol Cell* **36**, 340-347.
- Yamanaka, S., Poksay, K. S., Arnold, K. S. & Innerarity, T. L. (1997). A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. *Genes Dev* **11**, 321-333.
- Yan, Q., Moreland, R. J., Conaway, J. W. & Conaway, R. C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J Biol Chem* **274**, 35668-35675.
- Yanagida, N., Yoshida, S., Nazerian, K. & Lee, L. F. (1993). Nucleotide and predicted amino acid sequences of Marek's disease virus homologues of herpes simplex virus major tegument proteins. *J Gen Virol* **74** (Pt 9), 1837-1845.

- Yang, W., Chendrimada, T. P., Wang, Q., Higuchi, M., Seeburg, P. H., Shiekhattar, R. & Nishikura, K. (2006).** Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol* **13**, 13-21.
- Yao, Y., Zhao, Y., Smith, L. P., Lawrie, C. H., Saunders, N. J., Watson, M. & Nair, V. (2009a).** Differential expression of microRNAs in Marek's disease virus-transformed T-lymphoma cell lines. *J Gen Virol* **90**, 1551-1559.
- Yao, Y., Zhao, Y., Smith, L. P., Watson, M. & Nair, V. (2009b).** Novel microRNAs (miRNAs) encoded by herpesvirus of turkeys: evidence of miRNA evolution by duplication. *J Virol* **83**, 6969-6973.
- Yao, Y., Zhao, Y., Xu, H., Smith, L. P., Lawrie, C. H., Sewer, A., Zavolan, M. & Nair, V. (2007).** Marek's disease virus type 2 (MDV-2)-encoded microRNAs show no sequence conservation with those encoded by MDV-1. *J Virol* **81**, 7164-7170.
- Yao, Y., Zhao, Y., Xu, H., Smith, L. P., Lawrie, C. H., Watson, M. & Nair, V. (2008).** MicroRNA profile of Marek's disease virus-transformed T-cell line MSB-1: predominance of virus-encoded microRNAs. *J Virol* **82**, 4007-4015.
- Yekta, S., Shih, I. H. & Bartel, D. P. (2004).** MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-596.
- Yeom, K. H., Lee, Y., Han, J., Suh, M. R. & Kim, V. N. (2006).** Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Res* **34**, 4622-4629.
- Yin, Q., McBride, J., Fewell, C., Lacey, M., Wang, X., Lin, Z., Cameron, J. & Flemington, E. K. (2008).** MicroRNA-155 is an Epstein-Barr virus-induced gene that modulates Epstein-Barr virus-regulated gene expression pathways. *J Virol* **82**, 5295-5306.
- Yin, Q., Wang, X., Fewell, C., Cameron, J., Zhu, H., Baddoo, M., Lin, Z. & Flemington, E. K. (2010).** MicroRNA miR-155 inhibits bone morphogenetic protein (BMP) signaling and BMP-mediated Epstein-Barr virus reactivation. *J Virol* **84**, 6318-6327.
- Yu, X., Liu, L., Wu, L., Wang, L., Dong, C., Li, W. & Li, Q. (2010).** Herpes simplex virus type 1 tegument protein VP22 is capable of modulating the transcription of viral TK and gC genes via interaction with viral ICP0. *Biochimie* **92**, 1024-1030.
- Zahn, R. C., Schelp, I., Utermohlen, O. & von Laer, D. (2007).** A-to-G hypermutation in the genome of lymphocytic choriomeningitis virus. *J Virol* **81**, 457-464.
- Zawel, L., Kumar, K. P. & Reinberg, D. (1995).** Recycling of the general transcription factors during RNA polymerase II transcription. *Genes Dev* **9**, 1479-1490.
- Zeng, Y. & Cullen, B. R. (2004).** Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res* **32**, 4776-4785.
- Zhang, Y. & Rom, W. N. (1993).** Regulation of the interleukin-1 beta (IL-1 beta) gene by mycobacterial components and lipopolysaccharide is mediated by two nuclear factor-IL6 motifs. *Mol Cell Biol* **13**, 3831-3837.
- Zhang, Y., Sirko, D. A. & McKnight, J. L. (1991).** Role of herpes simplex virus type 1 UL46 and UL47 in alpha TIF-mediated transcriptional induction: characterization of three viral deletion mutants. *J Virol* **65**, 829-841.
- Zhao, Y., Kurian, D., Xu, H., Petherbridge, L., Smith, L. P., Hunt, L. & Nair, V. (2009a).** Interaction of Marek's disease virus oncoprotein Meq with heat-shock protein 70 in lymphoid tumour cells. *J Gen Virol* **90**, 2201-2208.
- Zhao, Y., Xu, H., Yao, Y., Smith, L. P., Kgosana, L., Green, J., Petherbridge, L., Baigent, S. J. & Nair, V. (2011).** Critical Role of the Virus-Encoded MicroRNA-155 Ortholog in the Induction of Marek's Disease Lymphomas. *PLoS Pathog* **7**, e1001305.
- Zhao, Y., Yao, Y., Xu, H., Lambeth, L., Smith, L. P., Kgosana, L., Wang, X. & Nair, V. (2009b).** A functional MicroRNA-155 ortholog encoded by the oncogenic Marek's disease virus. *J Virol* **83**, 489-492.
- Zhu, J., Mayeda, A. & Krainer, A. R. (2001).** Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell* **8**, 1351-1361.

PUBLICATIONS SCIENTIFIQUES

Marek's disease virus microRNA designated Mdv1-pre-miR-M4 targets both cellular and viral genes

Benoît Muylkens · Damien Coupeau ·
Ginette Dambrine · Sascha Trapp · Denis Rasschaert

Received: 24 February 2010 / Accepted: 26 July 2010
© Springer-Verlag 2010

Abstract Mdv1-miR-M4 is one of 25 microRNAs (miRNAs) expressed by Marek's disease virus (MDV-1), an oncogenic alphaherpesvirus that induces fatal T-cell lymphoma in chickens. Mdv1-miR-M4 was shown to be the second functional viral ortholog of miR-155, a cellular miRNA that plays a crucial role in several physiological and pathological processes in lymphocyte biology. In this study, we investigated a panel of putative mdv1-miR-M4 targets involved in gene networks affecting both cellular and viral life cycles. Using luciferase reporter assays, we showed that mdv1-miR-M4-5P and miR-155 efficiently targeted a common set of 3' untranslated regions (3'UTR) of six cellular genes (GPM6B, RREB1, c-Myb, MAP3-K7IP2, PU.1 and C/EBP). In addition, we also investigated the interactions between mdv1-miR-M4-5P and mdv1-miR-M43P and viral mRNAs encoding UL28 and UL32 in both reporter and western blot assays. Mdv1-miR-M4 specifically inhibited the translation of these two viral

proteins, which are involved in the cleavage/packaging of herpesvirus DNA.

Introduction

The discovery of miRNAs encoded by DNA viruses suggests that these viruses exploit the RNA silencing pathway to regulate host and viral genes. Most of the viruses with miRNAs belong to the herpesvirus family, and more than 140 herpesvirus miRNAs have been identified to date [1–4]. Members of the subfamilies *Alphaherpesvirinae*, *Betaherpesvirinae* and *Gammapherpesvirinae* have been shown to encode miRNAs. Marek's disease virus 1 (MDV-1), an alphaherpesvirus causing T-cell lymphoma in chickens, has been shown to express a panel of 25 MDV-1 miRNAs processed from 13 pre-miRNAs [5–9]. These pre-miRNAs are located within two genomic clusters mapping to the *meq* oncogene region and the LAT region in the terminal and inverted repeat long and short regions (IR_L, TR_L and IR_S, TR_S) of the MDV-1 genome [5].

The major characteristic common to all members of the herpesvirus family is an ability to establish and maintain latent infection during the lifetime of the virus in its host. Latency is associated with very limited virus gene expression. The tumorigenesis resulting from infection with human gammaherpesviruses, Epstein-Barr virus (EBV) or Kaposi's sarcoma human-associated virus (KSHV) and MDV-1 in chicken is associated with latency. Complex interactions take place between cells productively or latently infected with herpesvirus, neoplastically transformed cells and the immune system. The herpesvirus genome can subsequently be reactivated in latently infected cells. A balance is therefore established between latent infection and productive herpesvirus replication on the one

B. Muylkens and D. Coupeau contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-010-0777-y) contains supplementary material, which is available to authorized users.

B. Muylkens · D. Coupeau · G. Dambrine · S. Trapp ·
D. Rasschaert (✉)
Transcription, Lymphome Viro-Induit, University François
Rabelais, UFR Sciences et Techniques, Avenue Monge,
37200 Tours, France
e-mail: denis.rasschaert@univ-tours.fr

G. Dambrine
Department of Animal Health INRA, Nouzilly, France

B. Muylkens
Embryology, Veterinary Department, Faculty of Sciences,
University of Namur, FUNDP, 5000 Namur, Belgium

hand, and host immune surveillance and virus immune evasion on the other. Herpesvirus miRNAs are obvious candidates for involvement in the modulation of both of these phenomena, particularly as a significant proportion of herpesvirus genes have been shown to modulate host cell responses to virus infection [10]. Herpesvirus miRNAs have been shown to regulate the transition from latent to lytic replication by targeting immediate-early (IE) genes [4]. Herpesvirus miRNAs may also target viral genes expressed later, during viral replication, such as the genes encoding DNA polymerase [1, 11] and the viral uracil DNA glycosylase UL114 [12], thereby regulating levels of viral replication.

Other herpesvirus miRNAs may target host immune response genes, enabling the virus to escape immune recognition and preventing the destruction of herpesvirus-infected cells. HCMV miR-UL112-1 has been shown to target MICB, the major histocompatibility complex class I-related chain B, a ligand essential for the killing activities of NK cells [13], whereas EBV miR-BHRF1-3 inhibits the production of CXCL11 (CXC-chemokine ligand 11), a secreted T-cell attractant, thereby enabling virus-infected cells to evade the immune system [14]. The miRNAs encoded by oncogenic herpesviruses may also be more directly involved in tumorigenesis. Inhibition of the proapoptotic protein PUMA (p53-upregulated modulator of apoptosis) by EBV miR-BART5 and the targeting of thrombospondin 1 (THBS1) by KSHV miRNAs both enhance tumor survival [15, 16]. KSHV miR-K12-11 and mdv1-miR4-5P, two herpesvirus orthologs of cellular miR-155 with strictly identical seed sequences, may be particularly important, playing a multifunctional role in viral tumorigenesis [8, 17–19]. KSHV miR-K12-11 and hsa-miR-155 have overlapping sets of cellular targets, including BACH1 (BTB and CNC homology 1), a transcriptional repressor of heme oxygenase 1 (HMOX1), a protein that enhances cell survival and proliferation [17]. Furthermore, miR-155 has been shown to play a critical role in regulating lymphocyte hemostasis and immune tolerance [20–22]. It also has been identified as a component of the primary macrophage response to inflammatory mediators and as a protective miRNA controlling local tissue damage due to inflammation. The deregulation of miR-155 seems to play an important role in several viral infections. During the infection of humans with T-lymphotropic virus type 1 (HTLV-1) or chickens with reticuloendotheliosis virus strain T (REV-T), miR-155 has been shown to be overexpressed [23, 24]. The EBV immortalization of B lymphocytes has been shown to be associated with the induction of miR-155, which plays an important role in keeping the EBV genome stable during latent infection [25, 26]. By contrast, miR-155 has been shown to be downregulated in KSHV and MDV infections, although

a functional homolog of miR-155 is expressed during infections with these two viruses [8, 17–19, 27]. Further investigations are required to determine the effects of miR-155 deregulation and the role of functional orthologs in herpesvirus pathogenesis. MiR-155 is processed from a primary transcript, referred to as B-cell integration cluster (BIC) [28]. The *BIC* gene was originally identified as a frequent integration site of avian leukosis virus in lymphomas [29], and activation of the *BIC* gene can accelerate the pathogenesis of lymphomas and leukemia that are associated with the up-regulation of c-MYC [30]. Thus, miR-155 function in posttranscriptional down-regulation of gene expression could be responsible for the oncogenic activity attributed to *BIC* RNA.

In this study, we focused on mdv1-pre-miR-M4, which gives rise to mdv1-miR-M4-5P, an ortholog of miR-155. During the natural infection of chickens with MDV, we found that the two arms, mdv1-miR-M4-5P and mdv1-miR-M4-3P, of the pre-miR4 molecule were expressed, at times corresponding to latency and lymphomagenesis, but at different levels. We then investigated the functionality of these two miRNAs. We first showed that gga-miR-155 and mdv1-miR-M4-5P, whether expressed ectopically or endogenously, targeted a common set of cellular genes that are potentially involved in tumorigenesis. We experimentally validated three new targets of miR-155 and its orthologs, i.e. c-Myb, GPM6B and RREB1. We also identified viral targets for mdv1-miR-M4-5P and mdv1-miR-M4-3P, the viral UL28 and UL32 proteins, respectively. We found that, although gga-miR-155 and mdv1-miR-M4-5P had the same seed sequence (nucleotides 2–8 of the miR are perfectly complementary to the target sequence), only mdv1-miR-M4-5P targeted UL28. This is the first report to show that a miR-155 ortholog can target both cellular and viral proteins, demonstrating a multifunctional effect of miR-155 orthologs. These MDV-1 proteins are homologous to human herpesvirus 1 (HHV-1) proteins required for the cleavage and packaging of virion DNA. Their role in MDV-1 replication has yet to be investigated, but such studies could be based on the ectopic expression of mdv1-pre-miR-M4.

Materials and methods

miRNA and genome sequences

MDV-1 and chicken miRNA sequences were downloaded from the miRBase release 10.1 sequence database (<http://www.mirbase.org/>). The annotated genome sequence for MDV-1 strain MD5 was downloaded from NCBI (GenBank reference number AF243438). The chicken genome “assembly 2.1” (http://www.ensembl.org/Gallus_gallus) was

used to design primer sets for PCR that bound the predicted 3' UTRs.

Target prediction

The online target prediction algorithm Targetscan (<http://www.targetscan.org/>) was used to identify potential cellular target sites from the *Gallus gallus* 3' UTR database. Blast analysis of mdv1-miR-M4-5P and -3P sequences against the MDV-1 genome were carried out with the accessory application “local blast” available through BioEdit version 7.0.5 sequence alignment software.

Cell lines

The DF-1 chicken fibroblast cell line was cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (FCS) and 5% chicken serum. The LMH cell line, obtained from a chemically induced hepatocellular carcinoma, was cultured in gelatin-coated dishes. Two transformed lymphoblastoid cell lines, MDCC-MSB-1 and MDCC-54-O, were propagated in complete Roswell Park Memorial Institute medium (RPMI) supplemented with 10% fetal calf serum and 5% chicken serum. These two cell lines

are latently infected with MDV-1. The MSB-1 cell line, one of the first MDV-infected cell lines isolated [31], is derived from a spleen lymphoma induced by a virulent strain of MDV-1, and is currently used as a reference. It was previously shown to be co-infected with both MDV-1 (strain BC-1) and MDV-2 (strain HPRS24) [32]. The 54-O cell line was established at INRA, from a chicken ovary lymphoma induced by the very virulent MDV-1 strain RB-1B reconstituted from an infectious bacterial artificial chromosome clone (kindly provided by Venugopal Nair). The DT40 B-cell line, which is derived from an avian leukosis virus-induced lymphoma, was cultured at 41°C, under an atmosphere containing 5% CO₂, in DMEM supplemented with 10% FCS.

MiRNA expression constructs

The chicken miR-155 and miR-21 precursors were amplified from *Gallus gallus* genomic DNA (PA-12 white leghorn chicken line). The mdv1-miR-M4 and miR-M7 precursors were amplified from RB-1B genomic DNA. All the pre-miRNA sequences were amplified with adjoining sequences, using primer pairs (Table 1) with *NheI* and *HindIII* extensions, as described by Clop et al. [33]. The *NheI*/

Table 1 Oligonucleotides used in this study

Name	Nucleotide sequence 5' → 3'	Restriction site ^a
AS-mdv1-miR-M4-5P	ACGAAGGGTTCGATACAGCATTA	
AS-mdv1-miR-M4-3P	GGTCATGCTGTCAGAACCAT	
AS-gga-miR-155	CCCCTATCACGATTAGCATTA	
AS-gga-miR-21	TCAACATCAGTCTGATAAGCTA	
Fwd premir_155	TTTGCTAGCAGAGTTCTTCTGTAGGCTGT	<i>NheI</i>
Rev premir_155	TTTAAGCTTCACAGCAGCAGTTTGTTCCT	<i>HindIII</i>
Fwd premir_M4	TTTGCTAGCTGGGCTTGTTTGAATGTCC	<i>NheI</i>
Rev premir_M4	TTTAAGCTTTACGATGGTTTTCAGCGAT	<i>HindIII</i>
Fwd premir_21	TTTGCTAGCTGTTGCCAGCCGTCGTGACAT	<i>NheI</i>
Rev premir_21	TTTAAGCTTGCTGGAGATGGGTGAGCAA	<i>HindIII</i>
Fwd premir_M7	TTTGCTAGCGATGCTCTCTAGCCAAGAGAG	<i>NheI</i>
Rev premir_M7	TTTAAGCTTGCAGTTCTGAGGACACATTT	<i>HindIII</i>
CEBPβ fwd	CAGTCTAGACTTTAAGCATGGCTGGGAAT	<i>XbaI</i>
CEBPβ rev	CAGACTAGTAAAGTTTATTAAATCGTTC	<i>SpeI</i>
CEBPβ mut fwd	GCTACATTTGAATTTCGAAATGAAGCATTTTAATAAAC	
CEBPβ mut rev	AAATCGTTCATTTTCGAATTCAAATGTAGCAAGAATGC	
MAP3K7IP2 fwd	CAGTCTAGACACAGTGAAATACCCCAAAGC	<i>XbaI</i>
MAP3K7IP2 rev	CAGACTAGTTGATCTGTTGGTGGGTTGAA	<i>SpeI</i>
MAP3K7IP2 mut fwd	CCTTAAATAAGATTTCGAAATAAGAACTATGCACAAAATAGC	
MAP3K7IP2 mut rev	GTTTCTTATTTCGAATCTATTTAAGGCATTGCTTCTCTC	
GPM6B fwd	CAGTCTAGACCTGTTGCATGAAGGAGACA	<i>XbaI</i>
GPM6B rev	CAGACTAGTCAGCTACAGGTCTGTTTAA	<i>SpeI</i>

Table 1 continued

Name	Nucleotide sequence 5' → 3'	Restriction site ^a
GPM6B mut fwd	ATGTAAAGAT TCG AAAATTTTGCTATGTATAAAATTTTGTAACC	
GPM6B mut rev	CATAGCAAAATTT TCGA ATCTTTACATAAAAAATATTAGAC	
RREB1 fwd	CAGTCTAGAAGGTTTCCAGGGTCTCGTTT	<i>XbaI</i>
RREB1 rev	CAGACTAGTGAAATTCTTTGCGCTTCCTG	<i>SpeI</i>
RREB1 mut 1 fwd	CATTTAAT TCG AAATGTTACCTTTTGTATTGGTGTGTGTG	
RREB1 mut 1 rev	GGTAACATTT TCGA ATTTAAATGAAAAAGGTCCTCTCC	
RREB1 mut 2 fwd	CCACACGGTGGAT TCG AAAGCCTGTTCAGGCATTAGCAG	
RREB1 mut 2 rev	GGAACAGGCTTT TCGA ATCCACCGTGTGGCAAATGCGAG	
c-Myb fwd	CAGGCGGCCGCGACTTCTCCGGAGAAGC	<i>NotI</i>
c-Myb rev	CAGGCGGCCGCAATTAAGTGCAGGTGAAGC	<i>NotI</i>
c-Myb mut 1 fwd	CAGGCGGCCGCGACTTCTCCGGAGAA TCGA ATGGTGG	
c-Myb mut 2 fwd	GAAGTAAAAA TCGA ACTACACACGCTAGTCTGTGC	
c-Myb mut 2 rev	GTAGT TCGA ATTTTTTACTTCAGTATTTTACACCAATGC	
PU.1 fwd	CAGTCTAGAGCCACGGGACCCTAAGCAA	<i>XbaI</i>
PU.1 rev	CAGACTAGTAAATATGTACATGTTATTAC	<i>SpeI</i>
PU.1 mut 1 fwd	CAGTGACACT TCGA AGCTTACAGCTTTCAAAGTAGC	
PU.1 mut 1 rev	GCTGTAAAGCT TCGA AAAGTGTCACTGTGAGTGACACAGC	
PU.1 mut 2 fwd	GCTTTCAAAGT TCGA AAATACAGGCTCTGCCGAGCTC	
PU.1 mut 2 rev	GCCTGTATTT TCGA ATAGTTTGAAAGCTGTAAGCTAATGCAG	
PU.1 mut x2 fwd	CAGTGACACT TCGA AGCTTACAGCTTTCAAAGT TCGA AAATACAGGCTCTGCCGAGCTC	
PU.1 mut x2 rev	GCCTGTATTT TCGA ATAGTTTGAAAGCTGTAAGCT TCGA AAAGTGTCACTGTGAGTGACACAGC	
UL28target fwd	CAGTCTAGAGGCGAAGTATTGATAGACTC	<i>XbaI</i>
UL28target rev	CAGACTAGTTTCGGCACGACAGATGTGC	<i>SpeI</i>
UL28 mut fwd	CCAGCCGTGCAATAGTGAAGGATATTTTGTGTTGAAAATG	
UL28 mut rev	CATTTTCAACACTAAAATATCCTTCACTATTTCGACGGCTGG	
UL28ORF fwd	ATGTTGGGAATGTCTCATAACCGGTT	
UL28ORF rev	GATGGGGGCGTGGCTGTTGGTAATCT	
UL32target fwd	CTGTTCTAGATGCTTGGCGATCA	<i>XbaI</i>
UL32target rev	CAGACTAGTTGTCGCCGACTCCGTCCAAGT	<i>SpeI</i>
UL32 mut fwd	TAAAGCCCGTCTT TCG CAGTCACTAGCCGTTTGAAGGTCGCGTATG	
UL32 mut rev	ACGCGACCTTCAAA CGG CTAGTGACTGCGAAGACGGGCTTTATAT	
UL32ORF fwd	ATGGCCAACCGCCCTACAGAGTTGGC	
UL32ORF rev	CACGTAGACTCCTAATGTATGCTCGA	
Sequencing primers		
Forward-PrLTK	GCAAGAAGATGCACCTGATG	
Reverse-PrLTK	TCACTGCATTCTAGTTGTGG	
UL28mid	CAATGAGAGACACTACTACCGA	
UL32mid	GGGTGATGCCAAAGTCTTGTA	

The mutations of seed matches are indicated in bold

^a Endonuclease restriction sites are underlined

HindIII-cleaved PCR products were inserted between the *NheI* and *HindIII* sites of the pcDNA3.1(+) vector (Invitrogen).

Luciferase reporter constructs

The predicted 3' UTR targets were amplified by PCR (Promega) from *Gallus gallus* genomic DNA (PA-12 white leghorn chicken line). Viral CDS targets were amplified from RB-1B viral genomic DNA. Wild-type targets were amplified with the primer pairs indicated (Table 1), with *NotI* extensions for c-Myb and with *XbaI* and *SpeI* extensions for all other targets. The mutant complementary sites were generated by overlapping PCR mutagenesis, as described previously [34]. Wild-type or mutated PCR products were inserted into pGEMT Easy Vector (Promega), cleaved with *NotI* for c-Myb or with *XbaI/SpeI* for all other targets, and inserted into the 3' UTR of the *Renilla* luciferase gene of the pRL-TK vector (Promega), which had been cleaved with the appropriate restriction endonuclease. Plasmid DNAs were purified using a Nucleobond Xtra midi kit (Macherey–Nagel), and the inserts of all vectors were fully sequenced.

Expression constructs for viral proteins expression

The MDV1 UL28 and UL32 coding sequences were obtained by PCR amplification from RB-1B genomic DNA, using gene-specific primers (Table 1). The primers were designed to ensure that the stop codon was omitted, to allow expression of the epitope tag (V5&His). PCR products were inserted into the pcDNA3.1/V5&His TOPO cloning vector (Invitrogen), and plasmids were amplified according to the manufacturer's protocol.

Luciferase reporter assays

For reporter assays with miRNA expression vectors, DF-1 chicken fibroblast cells were cotransfected with 1.2 µg of miRNA expression vector, 60 ng of luciferase reporter constructs containing wild-type or mutant putative target sites and 15 ng of luciferase control vector, in 24-well plates (6×10^4 cells/well) using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocols. DT40 B-cells were cotransfected by resuspending 2×10^6 cells in Nucleofactor solution T (Amaxa Biosystems, Cologne, Germany) and electroporated using Nucleofactor program B-023. For all assays, 2×10^6 cells were electroporated in the presence of 4 µg of miRNA expression vector, 200 ng of luciferase reporter constructs containing wild-type or mutant putative miRNA response elements and 50 ng of luciferase control vector. For reporter assays with endogenously expressed viral miRNAs, two MDV-1

lymphoblastoid cell lines, MDCC-MSB-1 and MDCC-54-O, were electroporated using an Equibio "Easyject Plus" electroporator (single pulse, 400 V/mm, 1,500 µF). For all assays, 8×10^6 cells were electroporated in the presence of 40 µg of luciferase reporter constructs with wild-type or mutant putative target sites and 500 ng of luciferase control vector. Luciferase activity was quantified with the Dual-Luciferase Reporter Assay system (Promega), according to the manufacturer's protocol. For the standardization of inter- and intra-assay luciferase activity, we systematically used the control vector pcDNALuc, from which the firefly luciferase gene is expressed under the control of the CMV promoter. Firefly and *Renilla* luciferase activities were measured consecutively 24 h after transfection. Luminescence was measured with a luminometer (Mithras LB940 luminometer; Berthold Technologies). Three independent experiments were carried out in triplicate.

Detection of UL28 and UL32 V5-tagged proteins following transient miRNA expression

LMH cells, dispensed into 24-well plates at a density of 1.2×10^5 cells/well, were cotransfected with 0.2 µg of pcDNA3.1-UL28/V5 or pcDNA3.1-UL32/V5 and 1.8 µg of miRNA expression vector, in the presence of Lipofectamine 2000 (Invitrogen). Cells were treated with lysis buffer (15 mM Tris, 15 mM NaCl, 20 mM EDTA, 1% SDS, 5% beta-mercaptoethanol, 20% glycerol, bromophenol blue [pH 6.8]) 48 h after transfection. Proteins were separated by electrophoresis in 7.5% polyacrylamide gels containing SDS and electrophoretically blotted onto PVDF membranes. The membranes were blocked by incubation with 5% skim milk powder in PBST (PBS, 0.5% Tween 20) and probed with monoclonal mouse anti-V5 (Invitrogen) and monoclonal mouse anti-GAPDH (Millipore) primary antibodies, followed by polyclonal rabbit anti-mouse horseradish-peroxidase-conjugated secondary antibodies. The immunoreactions of interest were detected using a Femto Super Signal Western Detection Kit (Pierce Chemical Co.). GAPDH was used to demonstrate equal protein loading. Image J software (Wayne Rasband, NIH image, USA) was used to calculate the signal intensities for UL28 and UL32, relative to the GAPDH loading control.

Preparation of peripheral blood leukocytes from infected and mock-infected chickens

Four B13/B13 chickens were bred, raised and housed in isolated cages at the animal facility of the INRA-TOURS Research Center. Two chickens were used as mock-infected controls. Two chickens were inoculated, by the intramuscular route, at the age of 10 weeks, with 5×10^6 infectious peripheral blood leukocytes (PBLs) (wild-type

RB-1B). Before and 31 days after infection, PBLs were separated from whole blood samples by Ficoll-Hypaque gradient centrifugation as described previously [35].

RNA isolation and miRNA northern blotting analysis

RNA was extracted in TRIzol, and we then resolved 15 µg of total RNA per lane on a 15% acrylamide/bisacrylamide (19:1) gel containing 7 M urea, in Tris–borate-EDTA (TBE) buffer. Gels were stained briefly with ethidium bromide, and RNA was transferred onto a Nylon Plus membrane (Amersham) by electroblotting for 1.5 h at 350 mA in 0.50× TBE. The membrane was cross-linked by exposure to UV illumination (Stratalinker, Stratagene). The membrane was incubated with the prehybridization mixture for 1 h and then hybridized overnight in Perfect Hyb TM Plus hybridization buffer (Sigma) at 50°C, with 20 nM ³²P-5′ end-labeled DNA oligonucleotide probe complementary to the miRNA (Table 1). The membrane was then washed in low-stringency washing buffer. Blots were analyzed by phosphorimaging using a Storm 840 Molecular Dynamics machine.

Results

Expression profile of mdv1-miR-M4-5P, mdv1-miR-M4-3P and gga-miR-155 in MDV-1-infected peripheral blood leukocytes

Alignment of the mdv1-miR-M4-5P and gga-miR-155 sequences showed them to have identical seed sequences (nucleotides 1–8; Fig. 1b) [8, 19]. However, their 3′ ends differ in length and nucleotide composition, with only a stretch of four consecutive nucleotides (nt 10–13 in mdv1-miR-M4-5P; nt 11–14 in gga-miR-155) conserved between the two miRNAs. Northern blot analysis of total RNA extracted from uninfected cells (mock-infected PBLs and DT40) or from MDV-1-infected cells (infected PBLs and MDCC 54-O and MSB1 cell lines) showed that the orthologous miRNAs mdv1-miR-M4-5P and gga-miR-155 had different patterns of expression. Unlike gga-miR-21, which was detected in the five cellular RNA extracts, gga-miR-155 and mdv1-miR-M4-5P were detected only in uninfected chicken PBLs and MDV-1-infected cells, respectively (Fig. 1c). In the control DT-40 B-cell line, which is derived from an avian leukosis virus-induced bursal lymphoma, we detected neither mdv1-miR-M4-5P nor gga-miR-155. As reported in previous studies analyzing expression of the whole panel of MDV-1 miRNAs [5, 7–9], the mdv1-miR-M4-3P was detected by northern blotting in the two MDV-1-infected cell lines with high levels of MDV-1-miRNA. Small RNA cloning assays

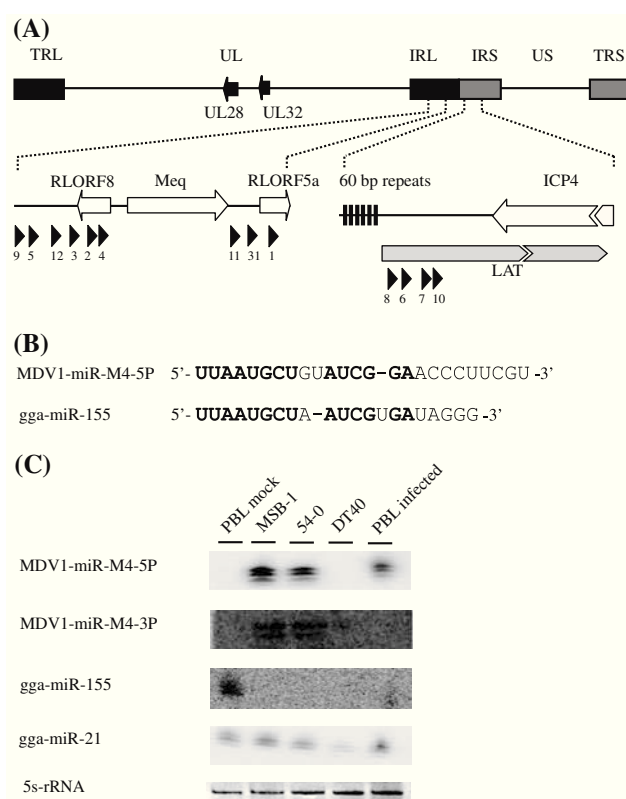


Fig. 1 Expression of mdv1-miR-M4-5P, mdv1-miR-M4-3P and gga-miR-155. **a** Localization of the 13 pre-miRNA loci in the MDV-1 genome. The MDV-1 genome is a class E genome including two unique sequences, a long sequence (UL) and a short sequence (US). These unique sequences are each flanked by two repeated and inverted sequences: internal repeat long (IRL) and terminal repeat long (TRL) surrounding UL; internal repeat short (IRS) and terminal repeat short (TRS) surrounding US. Cluster 1 exists in two inverted repeated copies within IRL and TRL and encompasses pre-miR-M9, -M5, -M12, -M3, -M2, -M4, -M11, -M31 and -M1. It extends up- and downstream from the ORF encoding the principal MDV-1 oncogene, *Meq*. Pre-miRNA-M2 and -M4 are antisense to the putative RLORF8 coding sequence. Pre-miRNA-M1 is located in the putative RLORF5a coding sequence. Cluster 2 exists as two inverted repeated copies within IRS and TRS and encompasses pre-miR-M8, -M6, -M7 and -M10. It is located at the start of the 10-kb latency-associated transcript. The location of the UL28 and UL32 coding sequences is indicated on the map of the MDV-1 genome. **b** Alignment of mature mdv1-miR-M4-5P and gga-miR-155. Nucleotides identical in the two miRNAs are indicated in *bold*. **c** Detection of mdv1-miR-M4-5P, gga-miR-155 and gga-miR-21. Northern blot analysis of total RNA (15 µg/lane) extracted from MSB-1, 54-O, DT40 cells and PBLs from MDV-1-infected chickens (PBL infected) and uninfected chickens (PBL mock). The membranes were hybridized with ³²P-labeled probes specific for mdv1-miR-M4-5P, mdv1-miR-M4-3P, gga-miR-155 and gga-miR-21. The 5S ribosomal RNA (stained with ethidium bromide, EB) was used to demonstrate equal RNA loading

demonstrated that both strands processed from pre-miR-4 were relatively abundant among total small RNA sequences. In sets of 1,425 and 1,885 small RNA clones analyzed from the 54-O and MSB-1 cell lines, respectively, mdv1-miR-M4-5P cloning frequencies accounted for 14.9 and

14.6% of all viral miRNAs clones, whereas mdv1-miR-M4-3P cloning frequencies accounted for 1.1 and 1.9% (our unpublished data).

Gga-miR-155 and mdv1-miR-M4-5P target a common set of cellular gene 3' UTRs

As mdv1-miR-M4-5P and gga-miR-155 had identical seed sequences, we hypothesized that these two miRNAs might regulate a common set of cellular genes. Therefore, using target prediction algorithms from Targetscan custom version 4.0 [36], we selected six candidate targets among the large set of miR-155 cellular potential targets. The

GPM6B, RREB1, c-Myb, MAP3K7IP2, PU.1 and C/EBP β (Fig. S1) candidates had high miRE prediction scores and were selected for study on the basis of their involvement in lymphomagenesis, innate response or oncogenesis. The last two of these targets have already been validated for both miRNAs [19], and MAP3K7IP2 has been validated for hsa-miR-155 [37]. We investigated the direct interaction between these potential cellular targets and mdv1-miR-M4 and gga-miR-155. We classified the putative target genes into two groups, as a function of the number of miREs (one or two) in the 3' UTR sequence (Figs. 2, 3). The full-length or partial 3' UTRs of each selected potential miRNA target and the corresponding miRE-mutated 3' UTRs were

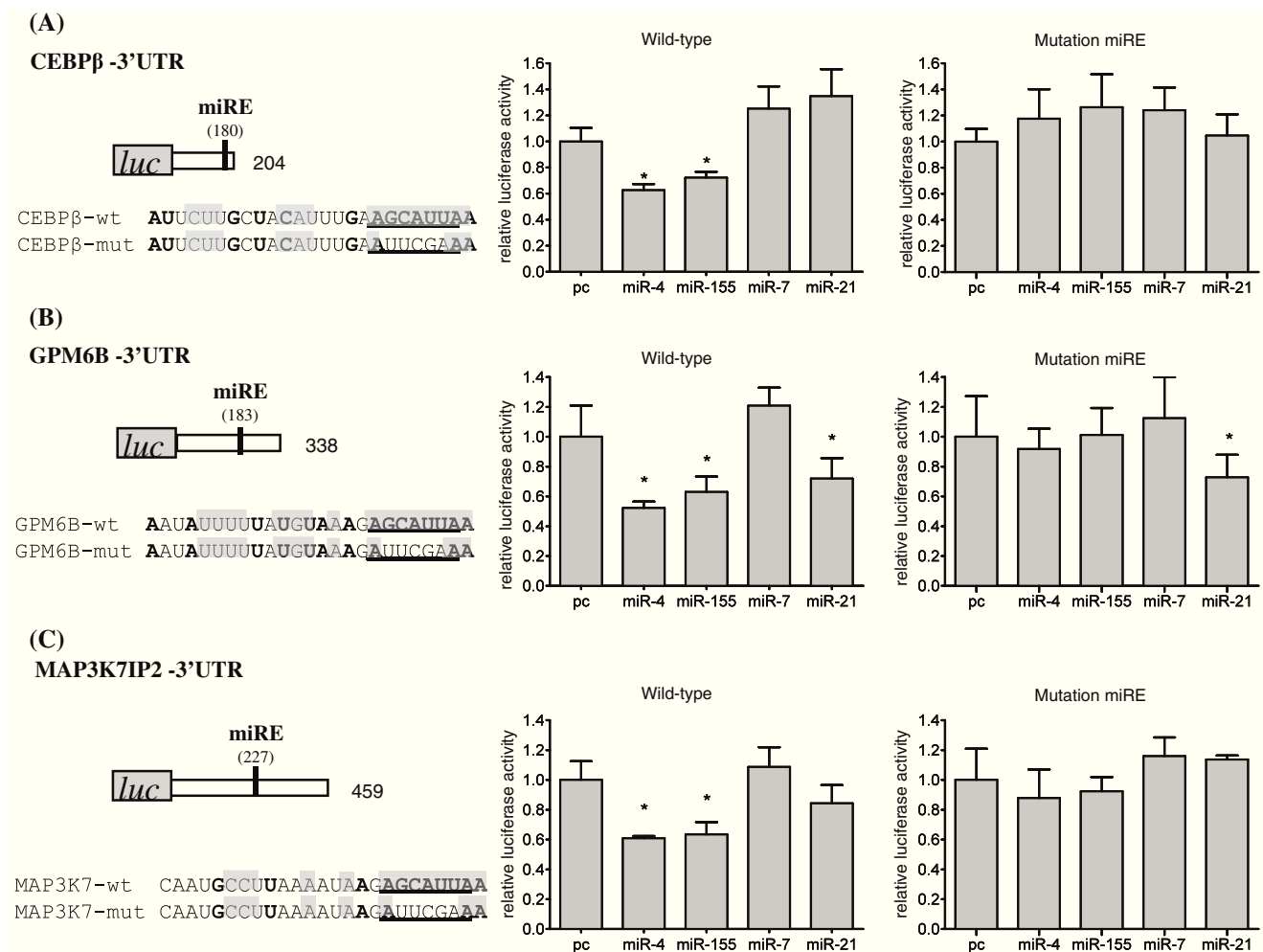


Fig. 2 Comparison of the efficacy of mdv1-miR-M4-5P and gga-miR-155 with three 3' UTRs, each bearing a single miRE. The 3' UTRs of *C/EBP β* (a), *GPM6B* (b) and *MAP3K7IP2* (c) carrying the predicted wild-type or respective mutant targets, were inserted downstream from the *Renilla* luciferase reporter gene in pRL-TK. Each construct was used to cotransfect DF-1 cells, together with pcDNA-mdv1-miR-M4, pcDNA-gga-miR-155, pcDNA-mdv1-miR-M7, pcDNA-gga-miR-21, or pcDNA. Reporter activities were standardized by measuring *Renilla*-to-firefly luminescence ratios. For each luciferase assay, relative luciferase activity was normalized

with respect to pcDNA3.1 empty vector. Error bars indicate the standard deviation (SD) for triplicate experiments. *Students *t* test *P* value <0.05 versus pcDNA3.1 empty vector. Schematic diagrams of the pRL-TK constructs are represented on the left. The vertical bar indicates the location of the miRE within the 3' UTR. The position of this first nucleotide is indicated in brackets. The wild-type and mutated seed match of targets is underlined. Target complementarities with mdv1-miR-M4-5P are indicated in bold, and target complementarities with gga-miR-155 are highlighted in gray

inserted downstream from the luciferase reporter gene in the PRL-TK vector. DF-1 cells were cotransfected with these reporter constructs and with pcDNA3.1 vectors expressing either mdv1-miR-M4 and gga-miR-155 or the control mdv1-miR-M7 and gga-miR-21. Similar effects were observed for all 3' UTRs with one putative miRE when mdv1-miR-M4 or gga-miR-155 were overexpressed (Fig. 2), with a larger ($P < 0.01$) decrease in luciferase activity for wild-type constructs than for the corresponding miRE-mutated constructs. No significant difference was found between wild-type 3' UTRs and the corresponding miRE-mutated constructs if co-transfections were carried out with an empty pcDNA vector or with the negative pcDNA control vectors expressing mdv1-miR-M7 or gga-miR-21. However, in comparison with the empty pcDNA vector, a significant effect of gga-miR-21 was observed using both wild-type and miRE-mutated GPM6B 3' UTR constructs, probably due to a gga-miR-21 potential response element present in the GPM6B 3' UTR sequence used in our assays that did not overlap with the specific mdv1-miR-M4 and gga-miR-155 miREs. Thus, the 3' UTRs of C/EBP β , MAP3K7IP2 and GPM6B are clearly functional targets of both mdv1-miR-M4 and gga-miR-155. For the other three genes, with 3' UTRs containing two potential miREs, the 3' UTRs of RREB1, c-Myb and PU.1 were functional targets of both mdv1-miR-M4 and gga-miR-155, as shown by the luciferase data obtained in comparisons of wild-type and double-mutated 3' UTRs. However, two different patterns were observed if the miREs were mutated separately. On the one hand, miRE1 and miRE2 within the 3' UTRs of both RREB1 and c-Myb (Fig. 3a, b, respectively) had potent but independent effects on functional targeting by mdv1-miR-M4 and gga-miR-155, with cooperation observed between the two target sites. However, comparison of the effects of mdv1-miR-M4 and gga-miR-155 on c-Myb miRE1 (Fig. 3b-mutation miRE1) showed that c-Myb miRE1 was targeted slightly more efficiently by the viral miRNA than it was by its cellular ortholog. On the other hand, only one of the two predicted miREs in the 3' UTR of PU.1 was functional. Indeed, the miRE1 mutation did not restore reporter activity, whereas the miRE2 mutation fully restored relative luciferase activity, demonstrating that miRE2 was the only functional site within the 3' UTR of PU.1 (Fig. 3c). These data demonstrate that the miREs detected in the 3' UTRs of six cellular genes (C/EBP β , MAP3K7IP2, GPM6B, RREB1, c-Myb and PU.1) are functional targets of mdv1-miR-M4 and gga-miR-155 in 3'UTR reporter assays. Except for c-myb miRE1, no obvious difference was detected in the levels of repression induced by cellular and viral miRNA.

For confirmation of the effects of endogenous mdv1-miR-M4-5P on the potential targets, the same constructs were tested in the MDV-1-transformed cell lines MSB-1

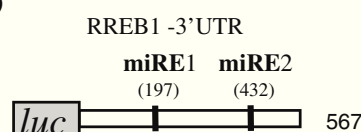
Fig. 3 Comparative efficacy of mdv1-miR-M4 and gga-miR-155 with three 3' UTRs, each bearing two miREs (miRE1 and miRE2). The 3' UTRs of *RREB1* (a), *c-Myb* (b) and *PU.1* (c) carrying the predicted wild-type or respective mutant targets, were inserted downstream from the *Renilla* luciferase reporter gene in pRL-TK. Each construct was used to cotransfect DF-1 cells, together with pcDNA-mdv1-miR-M4, pcDNA-gga-miR-155, pcDNA-mdv1-miR-M7, pcDNA-gga-miR-21, or pcDNA. Reporter activities were standardized by determining *Renilla*-to-firefly luminescence ratios. For each luciferase assay, relative luciferase activity was normalized with respect to pcDNA3.1 empty vector. Error bars indicate the standard deviation (SD) for triplicate experiments. *Students *t* test *P* value, <0.05 versus pcDNA3.1 empty vector. **Students *t* test *P* value, <0.001 versus pcDNA3.1 empty vector. Schematic diagrams of pRL-TK constructs are indicated on the left. The vertical bar indicates the location of the miREs within the 3' UTR. The position of the first nucleotide is indicated in brackets. For each target, the wild-type and mutated seed matches of targets are underlined. Target complementarities with mdv1-miR-M4-5P are shown in bold, and target complementarities with gga-miR-155 are highlighted in gray

and 54-O. These two cell lines produce large amounts of viral mdv1-miR-M4-5P, but no cellular gga-miR-155 (Fig. 1c). A similar effect was observed in the two lymphoblastoid cell lines, for all 3' UTRs containing a single miRE (C/EBP β , MAP3K7IP2, GPM6B) (Fig. 4a), and for the 3' UTRs containing two miREs (RREB1, c-Myb and PU.1; Fig. 4b). In particular, the two RREB1 miREs strongly inhibited reporter activity, with wild-type constructs displaying much lower levels of relative luciferase activity (80% in the 54-O cell line and 90% in the MSB-1 cell line) than the double miRE mutant construct. The inhibition yield attributed to miRE 1 or miRE 2 was much greater in the endogenous assay, in which miRE 2 displayed stronger inhibition than miRE 1. As described previously, the two miREs within the c-Myb 3' UTR inhibited reporter activity to similar extents, and miRE2 was the only functional target in the PU.1 3' UTR. Thus, the 3' UTRs of all the genes tested were targeted by endogenous mdv1-miR-M4-5P constitutively expressed by two different MDV-1-transformed cell lines, consistent with the data obtained for a transient non-lymphoid system.

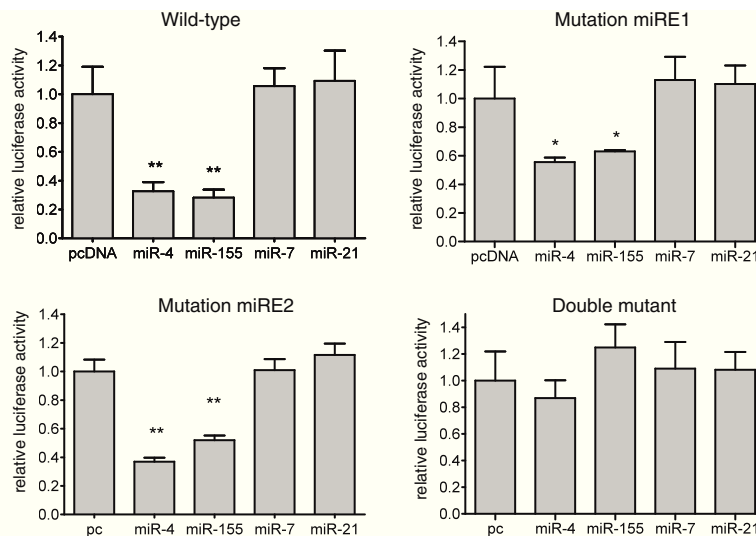
Mdv1-miR-M4 guide and passenger strands target viral genes involved in the cleavage/packaging of viral DNA

An internal basic local alignment search (BLAST) of mdv1-miR-M4-5P and mdv1-miR-M4-3P sequences in the MDV-1 genome showed antisense sequence similarities between mdv1-miR-M4-5P and the mRNA encoding UL28, and between mdv1-miR-M4-3P and the mRNA encoding UL32 (Fig. S1C). The two sites of potential interaction displayed extended matches beyond the miRNA seed region. Nucleotides 2–12 at the 5' end of mdv1-miR-M4-5P matched with an 11-nucleotide stretch at the 3' end of the UL28 coding region, whereas nucleotides 1–12 of

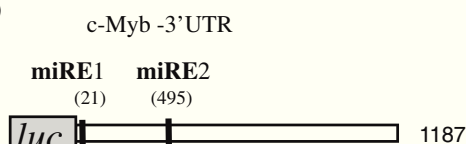
(A)



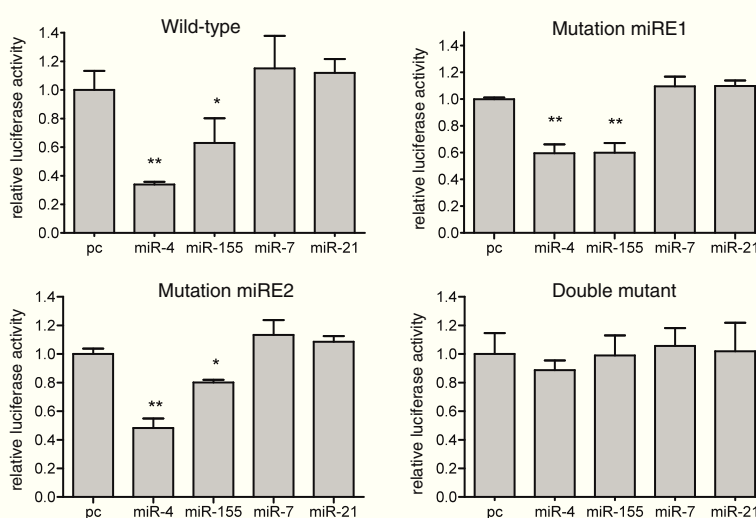
RREB1-S1 AGGACCUUUUC AUUUAAAGCAUUAA
RREB1-M1 AGGACCUUUUC AUUUAAAUUCGAAA
RREB1-S2 CAUUUGCCACACGGUGGAGCAUUAA
RREB1-M2 CAUUUGCCACACGGUGGAUUCGAAA



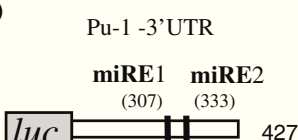
(B)



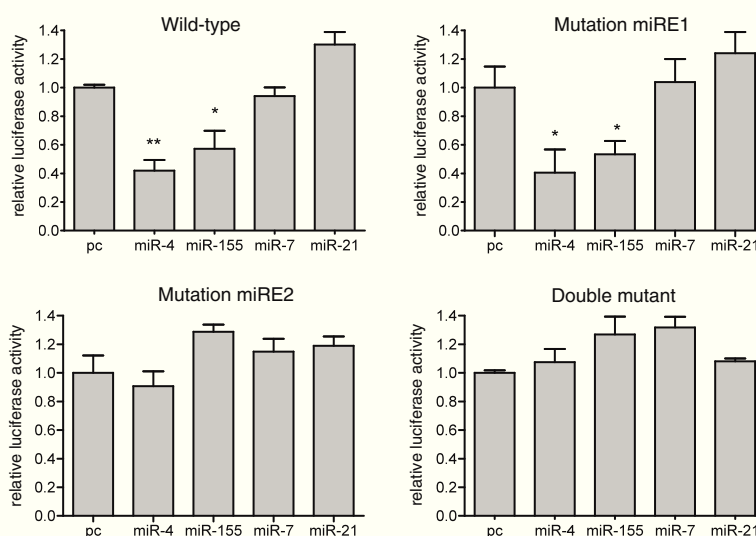
c-Myb-S1 CGCGACUUCUCCGGAGAGCAUUAAU
c-Myb-M1 CGCGACUUCUCCGGAGAAUUCGAAU
c-Myb-S2 AAAUACUGAAGUAAAAAGCAUUAAU
c-Myb-M2 AAAUACUGAAGUAAAAAUUCGAAU



(C)



Pu-1-S1 GUCACUCACAGUGACACUGCAUUAG
Pu-1-M1 GUCACUCACAGUGACACUUUCGAAAG
Pu-1-S2 CUUACAGCUUUCAAACUAGCAUUAA
Pu-1-M2 CUUACAGCUUUCAAACUAUUCGAAA



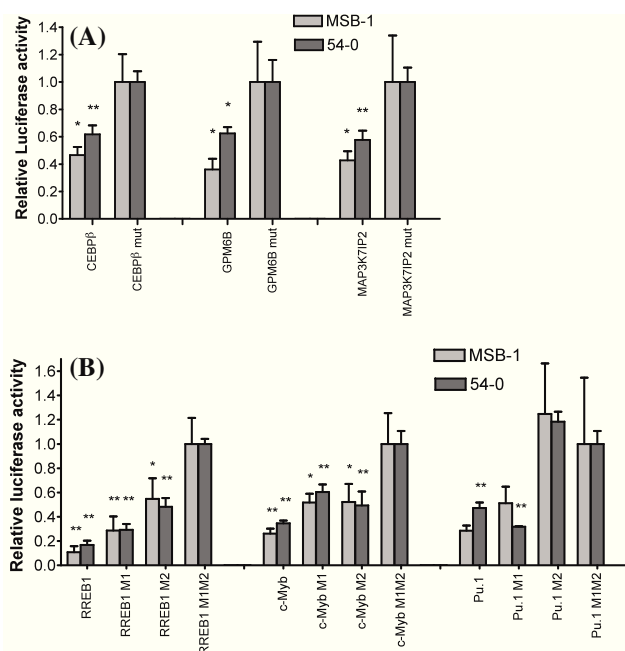


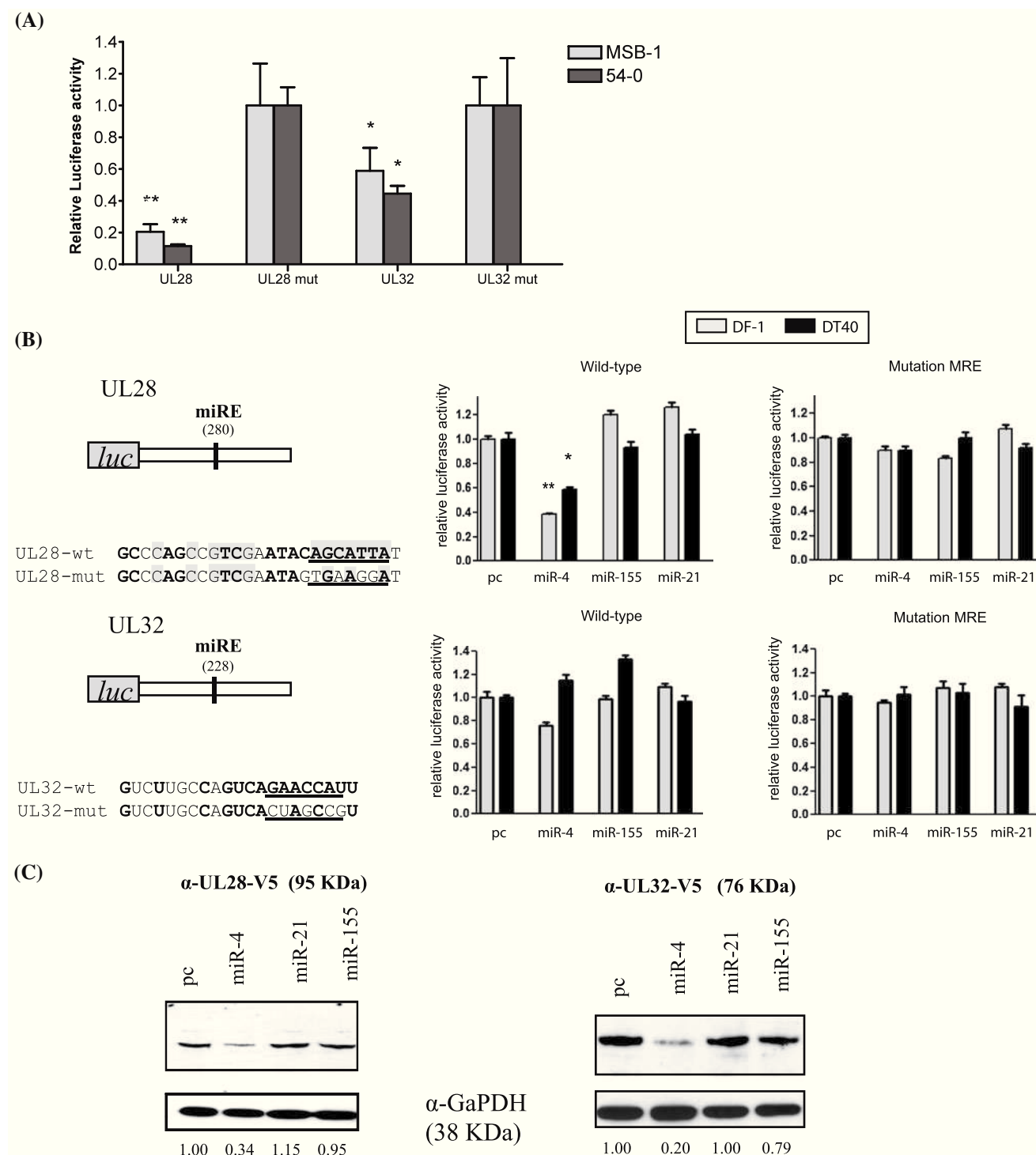
Fig. 4 Effect of endogenous mdv1-miR-M4 expression on the 3'UTR targets of chicken genes. Wild-type and corresponding mutated pRL-TK constructs bearing 3' UTR targets *C/EBPβ*, *GPM6B* and *MAP3K7IP2* (a) or *RREB1*, *c-Myb* and *PU.1* (b) were used to transfect MSB-1 and 54-O cell lines naturally expressing mdv1-miR-M4-5P. Reporter activities were standardized by determining *Renilla*-to-firefly luminescence ratios. Relative luciferase activity was normalized with respect to the reporter activity of the corresponding mutated target. Error bars indicate the standard deviation (SD) for triplicate experiments. *Students *t* test *P* value, <0.05 versus mutated target. **Students *t* test *P* value, <0.001 versus mutated target

mdv1-miR-M4-3P matched with a 12-nucleotide sequence at the 3' end of the UL32 coding region.

We investigated whether mdv1-miR-M4-5P and mdv1-miR-M4-3P affected the production of UL28 and UL32 proteins or mRNA, by first screening for potential miREs in a luciferase reporter assay, as described by Tay and coworkers, who investigated the miRNA-mediated regulation of the coding regions of *Nanog*, *Oct4* and *Sox2* [38]. The 500 nt surrounding the wild-type and mutated target sites of the UL28 and UL32 transcripts were inserted downstream from the luciferase reporter gene in pRL-TK vectors. These constructs were tested in two MDV-1-transformed cell lines constitutively expressing mdv1-miR-M4-5P and mdv1-miR-M4-3P, as demonstrated in Fig. 1 and in other reports [5, 7, 9] (Fig. 5a). The wild-type UL28 construct displayed much lower levels of luciferase activity (80 and 90% in the MSB-1 and 54-O cell lines, respectively) than the mutated construct. The decrease in luciferase activity was less marked for the UL32 constructs but was nonetheless similar to that observed when cellular targets with one miRE were tested in the same cells (Fig. 4a). We further investigated the interaction between

Fig. 5 The mdv1-miR-M4-5P and mdv1-miR-M4-3P target two viral genes. **a** Effect of endogenous mdv1-miR-M4 expression on the UL28 and UL32 viral targets. The wild-type and mutated predicted viral targets, surrounded by 500 bp of flanking sequence, were inserted into the plasmid downstream from the *Renilla* luciferase reporter gene. Each construct was used to transfect MSB-1 and 54-O cell lines naturally expressing mdv1-miR-M4-5P and -3P. Reporter activities were standardized by determining *Renilla*-to-firefly luminescence ratios. Relative luciferase activity was normalized with respect to the reporter activity of the mutated target. Error bars indicate the standard deviation (SD) for triplicate experiments. *Students *t* test *P* value, <0.05 versus mutated target. **Students *t* test *P* value, <0.001 versus mutated target. **b** Comparative targeting of mdv1-miR-M4 and gga-miR-155 in the UL28 and UL32 luciferase reporter assays. The wild-type and mutated predicted viral targets constructs were used to cotransfect DF-1 and DT40 cells, together with effector plasmids pcDNA, pcDNA-mdv1-miR-M4, pcDNA-gga-miR-155 and pcDNA-gga-miR-21. Reporter activities were normalized by determining *Renilla*-to-firefly luminescence ratios. For each luciferase assay, relative luciferase activity was normalized with respect to pcDNA3.1 empty vector. Error bars indicate the standard deviation (SD) of triplicate experiments. *Students *t* test *P* value <0.05 versus pcDNA3.1 empty vector. Schematic diagrams of the pRL-TK constructs are represented on the left. The vertical bar indicates the location of the miRE within the viral coding sequence, and the position of the first nucleotide is indicated in brackets. The wild-type and mutated seed matches of targets are underlined. For UL28 (a), target complementarities with mdv1-miR-M4-5P are shown in bold, and target complementarities with gga-miR-155 are highlighted in gray. For UL32 (b), target complementarities with mdv1-miR-M4-3P are shown in bold. (c) Western blot analysis of UL28, UL32 and chicken GAPDH levels. LMH cells were cotransfected with pcDNA3.1-UL28/V5 or pcDNA3.1-UL32/V5 and miRNA-expressing vectors pcDNA, pcDNA-mdv1-miR-M4, pcDNA-gga-miR-21 or pcDNA-gga-miR-155. Each blot was probed with mouse anti-V5 and mouse anti-GAPDH primary antibodies, followed by rabbit anti-mouse horseradish-conjugated secondary antibody. GAPDH was used to check for equal protein loading. Relative signal intensities for UL28 and UL32 were calculated with respect to the GAPDH loading control using Image J software (Wayne Rasband, NIH image, USA). The reported values (at the base of the blots) are the signals for UL28 and UL32 measured for miRNA-expressing vectors relative to the signals for UL28 and UL32 measured for the control vector expressing no miRNA. Similar results were obtained in three independent cotransfection assays. Data from a representative assay are presented

the two candidate target sites and the two mature strands generated from mdv1-pre-miR-M4 by evaluating reporter constructs upon transient expression of the mdv1-pre-miR-M4 and control pre-miRNAs in DF-1 and DT40 cells (Fig. 5b). The UL28 wild-type construct was associated with a highly significant decrease in luciferase activity in the DF-1 cell line, and a significant decrease in the DT40 cell line only when used to cotransfect cells with the mdv1-pre-miR-M4 expression vector. By contrast, no difference was observed between the wild-type UL28 construct and the corresponding miRE-mutated constructs if cotransfection was carried out with an empty pcDNA vector or the pcDNA control vectors. The UL32 wild-type construct was associated with a slight decrease (30%) in luciferase activity in the DF-1 cell line only if used for cotransfection



together with the pre-miR-M4 expression vector. However no decrease in luciferase activity was observed using UL32 wild-type construct in the DT40 cell line (Fig. 5b).

We then investigated the translational inhibition of UL28 and UL32 by the two mature strands generated from mdv1-pre-miR-M4, at the protein level. No antibodies were available against the UL28 and UL32 proteins of MDV-1.

We therefore first cloned and expressed an epitope-tagged version of each gene to generate proteins that could be detected in western blot experiments (Fig. 5c). Protein production was assessed in cotransfection experiments with pre-miRNA expression vectors and vectors encoding the epitope-tagged proteins. Mdv1-pre-miR-M4 decreased UL28 production by a factor of three and UL32 production

by 80% (Fig. 5c). Conversely, the production of UL28 and UL32 was not significantly decreased by coexpression with control pcDNA, gga-pre-miR-21 or orthologous gga-pre-miR-155. Thus, mdv1-pre-miR-M4 efficiently inhibits the production of the UL28 and UL32 proteins, whereas gga-pre-miR-155 does not.

Discussion

We focus here on the MDV microRNA mdv1-pre-miR-M4, containing the 5' arm of the RNA duplex mdv1-miR-M4-5P, which was previously identified as an ortholog of the cellular miR-155 [8, 19, 27]. We made use of the MDV-1 model, the only natural animal herpesvirus-induced tumor model, to assess the expression profile of mdv1-pre-miR-M4 in PBLs from chickens infected with the very virulent RB-1B strain of MDV-1, focusing on the time points corresponding to latency and lymphomagenesis. Several studies showing a differential expression of miRNAs between primary tissues and cell lines have highlighted the importance of analyzing miRNA expression in primary tissues [39–44]. We observed no expression of *Gallus gallus* gga-miR-155, whereas its ortholog, mdv1-miR-M4-5P, was abundantly expressed in PBLs from RB-1B-infected chickens at the time of lymphoma development, with mdv1-miR-M4-3P, the second strand of the miRNA duplex generated from mdv1-pre-miR-M4, also expressed, but to a lower level. Our results therefore extend the previously reported observations of gga-miR-155 downregulation and expression of both strands of mdv1-pre-miR-M4 during in vivo MDV pathogenesis previously reported for CD4+T-cell lines propagated in vitro from MDV-1-induced lymphoma [27]. Moreover, the relative expression levels of mdv1-miR-M4-5P and mdv1-miR-M4-3P were found to be similar in RB-1B-infected PBLs and in MDV-1 cell lines propagated in vitro. We also showed that both strands of mdv1-miR-M4 are functional, because mdv1-miR-M4-3P, although expressed less strongly than mdv1-miR-M4-5P, was nonetheless shown to target and repress UL32 production (Fig. 5b, c). These data are consistent with recent findings in *Drosophila melanogaster* showing that one strand of the RNA duplex, the guide strand, is preferentially selected and loaded into a silencing complex (RISC), but that several passenger strands may also be loaded, some of which have been demonstrated to be *bona fide trans*-regulatory RNAs with measurable effects on endogenous protein production [45–49].

In addition to identifying new genes targeted by both miR-155 and mdv1-miR-M4-5P, we showed that mdv1-miR-M4-5P, but not gga-miR-155, efficiently inhibited the production of UL28, a viral protein homologous to an HHV-1 protein involved in the cleavage/packaging of

virion DNA. Moreover, the less abundant strand processed from the 3' arm of the duplex, mdv1-miR-M4-3P, inhibited the production of another viral protein, UL32. The HHV-1 homolog of UL32 is also involved in the cleavage/packaging of virion DNA. Both strands of mdv1-miR-M4 downregulate viral protein production efficiently, via targets located within the protein coding regions, as recently reported for certain metazoans, not targeting the 3' untranslated region (UTR) of mRNAs [38, 50–52]. Moreover, mdv1-miR-M4-5P and gga-miR-155 had different inhibitory effects on UL28. Indeed, mdv1-miR-M4-5P strongly repressed UL28, whereas gga-miR-155 did not (Fig. 5b). The precise nucleotide requirements for the functional binding of a miRNA to a target sequence are not fully understood. Mdv1-miR-M4-5P displays 2–12 Watson–Crick base-pairings with the UL28 mRNA, corresponding to four base-pairings in addition to those of the seed sequence, whereas gga-miR-155 displays only canonical 7mer-m8 base-pairing, involving exclusively the seed region (Fig. S1). Thus, the additional 5' pairing between mdv1-miR-M4-5P and UL28 mRNA, enhancing binding specificity and the affinity of miRNA-target binding, may account for the differences observed between these two orthologs. The lack of functionality of gga-miR-155 despite a perfect seed pairing with UL28 mRNA is consistent with data observed for the *lcy-6* miRNA of *Caenorhabditis elegans* [53]. To identify the determinants that are specifically involved in mdv1-miR-M4/UL28 targeting, hybrids between mdv1-miR-M4 and gga-miR-155 were constructed, but these were not definitely informative (data not shown). Critical determinants involved in UL28 targeting remain to be identified and may likely act in a combinatorial manner [54]. Most of the still limited number of miRNA viral targets that have been validated experimentally display more extensive pairing with miRNA than cellular targets. The reasons for this stronger binding remain unclear.

By reporter luciferase assays in cells transiently expressing mdv1-pre-miR-M4 and gga-pre-miR-155, we showed that mdv1-miR-M4-5P functioned as a miR-155 ortholog with three previously characterized (C/EBP β , MAP3K7IP2 and PU.1) and three newly predicted (GPM6B, RREB1 and c-Myb) targets. For all six targets, no significant differences were observed between the levels of downregulation mediated by viral and cellular miRNAs, and the range of decrease observed was consistent with the published data for C/EBP β , MAP3K7IP2 and PU.1 [18, 19, 37, 55]. We also showed, in two MDV-1 lymphoblastoid cell lines, that the tested 3' UTRs were effectively targeted in the natural context of viral miRNA expression, using the same miREs as in transient experiments (Fig. 4). In addition to increasing the list of 3'UTRs targeted by mdv1-miR-M4-5P and gga-miR-155, the identification of three

new targets (GPM6B, RREB1 and c-Myb) provides additional evidence for an impact of miR-155 and its orthologs on pathways regulating lymphocyte differentiation and activation, and immune tolerance. Such regulation is important in the context of MDV. Indeed, the functional activity of monocytes and macrophages and nitric-oxide (NO) responses are critical in the control of MDV infection [56]. Mdv1-miR-4-5P may regulate macrophage activation and NO production through the same targets as hsa-miR-155 in human carcinoma [57] and mu-miR-155 in mouse [58].

Focusing on the infection cycle of MDV, one major consequence of cytokine regulation by mdv1-miR-M4-5P would be to regulate the balance between latent and lytic infection, which is known to be controlled by cytokine production [59]. This overall effect may be reinforced by the action of both the strands processed from the pre-miR-4 duplex, targeting two different proteins, the homologs of which in HHV1 are involved in a key step late in the viral replication cycle. We did not consider functional consequences for the viral cycle directly, but our data clearly demonstrate the inhibition of UL28 and UL32 protein production by the two strands processed from pre-miR-M4. UL28 and UL32 homologs have been found in all sub-families of mammalian and avian herpesviruses. However, the role of UL28 and UL32 in MDV packaging has yet to be demonstrated, and little is currently known about MDV packaging. Nonetheless, it appears likely that UL28 and UL32 are involved in the later stages of MDV replication. It is therefore possible that MDV1-miR-helps to maintain MDV-1 latency by downregulating the production of UL28 and UL32 and impairing late MDV morphogenesis and reactivation. This may constitute an additional mechanism used by herpesviruses to prevent inappropriate reactivation, in addition to those already described involving the repression of immediate-early viral genes, such as ICP0, ICP4, ICP34.5 and IE72, which are targeted by miRNAs from HHV-1, HHV-2 or HCMV [25, 52, 60].

In conclusion, our results demonstrate that the mdv1-premiR-M4, which is strongly expressed at the time of lymphomagenesis in MDV-1-infected chicken PBLs and processed into two mature and functional strands, is multifunctional. By extending the common set of experimentally validated 3' UTR cellular targets of the guide strand (mdv1-miR-M4-5P) and its previously described ortholog miR-155, we provide additional evidence for the probable involvement of mdv1-miR-M4-5P in modulating cytokine production and the chicken immune response to MDV infection. In particular, we have shown that mdv1-miR-M4-5P and mdv1-miR-M4-3P inhibit the production of the UL28 and UL32 proteins, respectively, thus providing the first evidence of late viral gene targeting by herpes miRNA and adding two new viral targets to the modest number of viral miRNA targets experimentally validated to date.

Homologs of both of these viral genes are involved in late HHV1 morphogenesis. Mdv1-pre-miR-M4 may therefore be a suitable tool for studies aiming to determine the precise role of UL28 and UL32 in the cleavage and packaging of viral DNA and late MDV morphogenesis, about which little is currently known.

Acknowledgments We thank Sébastien Pfeffer (IBMP, Strasbourg) for his advice concerning northern blots and small RNA cloning experiments and Sylvie Laurent for her advice concerning western blots. We thank Dr. Nair Venugopal (IAH, Compton) for providing the infectious BAC clone p-RB1B. Benoît Muyllens is a postdoctoral research fellow supported by the *Fonds National de la Recherche Scientifique (FNRS)*. This work was supported by the *Ligue Nationale contre le cancer, Comité du Cher*, AIP P0258 of the *Institut National de la Recherche Agronomique* and the *Agence Nationale de la Recherche (ANR-07-MIME-012-01)*.

References

- Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T (2004) Identification of virus-encoded micromRNAs. *Science* 304(5671):734–736
- Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser FA, van Dyk LF, Ho CK, Shuman S, Chien M, Russo JJ, Ju J, Randall G, Lindenbach BD, Rice CM, Simon V, Ho DD, Zavolan M, Tuschl T (2005) Identification of micromRNAs of the herpesvirus family. *Nat Methods* 2(4):269–276
- Gottwein E, Cullen BR (2008) Viral and cellular micromRNAs as determinants of viral pathogenesis and immunity. *Cell Host Microbe* 3(6):375–387
- Boss IW, Plaisance KB, Renne R (2009) Role of virus-encoded micromRNAs in herpesvirus biology. *Trends Microbiol* 17(12):544–553
- Burnside J, Bernberg E, Anderson A, Lu C, Meyers BC, Green PJ, Jain N, Isaacs G, Morgan RW (2006) Marek's disease virus encodes micromRNAs that map to meq and the latency-associated transcript. *J Virol* 80(17):8778–8786
- Burnside J, Morgan RW (2007) Genomics and marek's disease virus. *Cytogenet Genome Res* 117(1–4):376–387
- Burnside J, Ouyang M, Anderson A, Bernberg E, Lu C, Meyers BC, Green PJ, Markis M, Isaacs G, Huang E, Morgan RW (2008) Deep sequencing of chicken micromRNAs. *BMC Genomics* 9:185
- Morgan R, Anderson A, Bernberg E, Kamboj S, Huang E, Lagasse G, Isaacs G, Parcells M, Meyers BC, Green PJ, Burnside J (2008) Sequence conservation and differential expression of marek's disease virus micromRNAs. *J Virol* 82(24):12213–12220
- Yao Y, Zhao Y, Xu H, Smith LP, Lawrie CH, Watson M, Nair V (2008) MicromRNA profile of marek's disease virus-transformed t-cell line msb-1: predominance of virus-encoded micromRNAs. *J Virol* 82(8):4007–4015
- Areste C, Blackburn DJ (2009) Modulation of the immune system by kaposi's sarcoma-associated herpesvirus. *Trends Microbiol* 17(3):119–129
- Barth S, Pfuhl T, Mamiani A, Ehses C, Roemer K, Kremmer E, Jaker C, Hock J, Meister G, Grasser FA (2008) Epstein-barr virus-encoded micromRNA mir-bart2 down-regulates the viral DNA polymerase balf5. *Nucleic Acids Res* 36(2):666–675
- Stern-Ginossar N, Saleh N, Goldberg MD, Prichard M, Wolf DG, Mandelboim O (2009) Analysis of human cytomegalovirus-encoded micromRNA activity during infection. *J Virol* 83(20):10684–10693

13. Stern-Ginossar N, Elefant N, Zimmermann A, Wolf DG, Saleh N, Biton M, Horwitz E, Prokocimer Z, Prichard M, Hahn G, Goldman-Wohl D, Greenfield C, Yagel S, Hengel H, Altuvia Y, Margalit H, Mandelboim O (2007) Host immune system gene targeting by a viral mirna. *Science* 317(5836):376–381
14. Xia T, O'Hara A, Araujo I, Barreto J, Carvalho E, Sapucaia JB, Ramos JC, Luz E, Pedrosa C, Manrique M, Toomey NL, Brites C, Dittmer DP, Harrington WJ Jr (2008) Ebv micrnas in primary lymphomas and targeting of cxcl-11 by ebv-mir-bhrf1-3. *Cancer Res* 68(5):1436–1442
15. Choy EY, Siu KL, Kok KH, Lung RW, Tsang CM, To KF, Kwong DL, Tsao SW, Jin DY (2008) An epstein-barr virus-encoded microrna targets puma to promote host cell survival. *J Exp Med* 205(11):2551–2560
16. Samols MA, Skalsky RL, Maldonado AM, Riva A, Lopez MC, Baker HV, Renne R (2007) Identification of cellular genes targeted by kshv-encoded micrnas. *PLoS Pathog* 3(5):e65
17. Skalsky RL, Samols MA, Plaisance KB, Boss IW, Riva A, Lopez MC, Baker HV, Renne R (2007) Kaposi's sarcoma-associated herpesvirus encodes an ortholog of mir-155. *J Virol* 81(23):12836–12845
18. Gottwein E, Mukherjee N, Sachse C, Frenzel C, Majoros WH, Chi JT, Braich R, Manoharan M, Soutschek J, Ohler U, Cullen BR (2007) A viral microrna functions as an orthologue of cellular mir-155. *Nature* 450(7172):1096–1099
19. Zhao Y, Yao Y, Xu H, Lambeth L, Smith LP, Kgosana L, Wang X, Nair V (2009) A functional microrna-155 ortholog encoded by the oncogenic marek's disease virus. *J Virol* 83(1):489–492
20. Tili E, Croce CM, Michaille JJ (2009) Mir-155: on the crosstalk between inflammation and cancer. *Int Rev Immunol* 28(5):264–284
21. Faraoni I, Antonetti FR, Cardone J, Bonmassar E (2009) Mir-155 gene: a typical multifunctional microrna. *Biochim Biophys Acta*
22. Teng G, Papavasiliou FN (2008) Shhh! silencing by microrna-155. *Philos Trans R Soc Lond B Biol Sci* 364(1517):631–637
23. Pichler K, Schneider G, Grassmann R (2008) Microrna mir-146a and further oncogenesis-related cellular micrnas are dysregulated in htlv-1-transformed t lymphocytes. *Retrovirology* 5:100
24. Bolisetty MT, Dy G, Tam W, Beemon KL (2009) Reticuloendotheliosis virus strain t induces mir-155, which targets jarid2 and promotes cell survival. *J Virol* 83(23):12009–12017
25. Lu F, Weidmer A, Liu CG, Volinia S, Croce CM, Lieberman PM (2008) Epstein-barr virus-induced mir-155 attenuates nf-kappab signaling and stabilizes latent virus persistence. *J Virol* 82(21):10436–10443
26. Yin Q, Wang X, Fewell C, Cameron J, Zhu H, Baddoo M, Lin Z, Flemington EK (2010) Microrna mir-155 inhibits bone morphogenetic protein (bmp) signaling and bmp-mediated epstein-barr virus reactivation. *J Virol* 84(13):6318–6327. doi:[JVI.00635-10](https://doi.org/10.1128/JVI.00635-10)[pii][10.1128/JVI.00635-10](https://doi.org/10.1128/JVI.00635-10)
27. Yao Y, Zhao Y, Smith LP, Lawrie CH, Saunders NJ, Watson M, Nair V (2009) Differential expression of micrnas in marek's disease virus-transformed t-lymphoma cell lines. *J Gen Virol* 90(Pt 7):1551–1559
28. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T (2002) Identification of tissue-specific micrnas from mouse. *Curr Biol* 12(9):735–739. doi:[S0960982202008096](https://doi.org/10.1016/S0960982202008096)[pii]
29. Tam W, Ben-Yehuda D, Hayward WS (1997) Bic, a novel gene activated by proviral insertions in avian leukosis virus-induced lymphomas, is likely to function through its noncoding rna. *Mol Cell Biol* 17(3):1490–1502
30. Tam W, Hughes SH, Hayward WS, Besmer P (2002) Avian bic, a gene isolated from a common retroviral site in avian leukosis virus-induced lymphomas that encodes a noncoding rna, cooperates with c-myc in lymphomagenesis and erythroleukemogenesis. *J Virol* 76(9):4275–4286
31. Akiyama Y, Kato S, Iwa N (1973) Continuous cell culture from lymphoma of marek's disease. *Biken J* 16(4):177–179
32. Hirai K, Yamada M, Arao Y, Kato S, Nii S (1990) Replicating marek's disease virus (mdv) serotype 2 DNA with inserted mdv serotype 1 DNA sequences in a marek's disease lymphoblastoid cell line msb1-41c. *Arch Virol* 114(3–4):153–165
33. Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibe B, Bouix J, Caiment F, Elsen JM, Eycheenne F, Larzul C, Laville E, Meish F, Milenkovic D, Tobin J, Charlier C, Georges M (2006) A mutation creating a potential illegitimate microrna target site in the myostatin gene affects muscularity in sheep. *Nat Genet* 38(7):813–818
34. Fragnet L, Kut E, Rasschaert D (2005) Comparative functional study of the viral telomerase rna based on natural mutations. *J Biol Chem* 280(25):23502–23515
35. Djeraba-AitLounis A, Soubieux D, Klapper W, Rasschaert D (2004) Induction of telomerase activity in avian lymphoblastoid cell line transformed by marek's disease virus, mdcc-msb1. *Vet Pathol* 41(4):405–407
36. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microrna targets. *Cell* 120(1):15–20
37. Ceppi M, Pereira PM, Dunand-Sauthier I, Barras E, Reith W, Santos MA, Pierre P (2009) Microrna-155 modulates the interleukin-1 signaling pathway in activated human monocyte-derived dendritic cells. *Proc Natl Acad Sci USA* 106(8):2735–2740
38. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I (2008) Micrnas to nanog, oct4 and sox2 coding regions modulate embryonic stem cell differentiation. *Nature* 455(7216):1124–1128
39. Yeung ML, Yasunaga J, Bennasser Y, Duseti N, Harris D, Ahmad N, Matsuoka M, Jeang KT (2008) Roles for micrnas, mir-93 and mir-130b, and tumor protein 53-induced nuclear protein 1 tumor suppressor in cell growth dysregulation by human t-cell lymphotropic virus 1. *Cancer Res* 68(21):8976–8985
40. O'Hara AJ, Chugh P, Wang L, Netto EM, Luz E, Harrington WJ, Dezube BJ, Damania B, Dittmer DP (2009) Pre-microrna signatures delineate stages of endothelial cell transformation in kaposi sarcoma. *PLoS Pathog* 5(4):e1000389
41. Umbach JL, Cullen BR (2009) In-depth analysis of kaposi's sarcoma-associated herpesvirus microrna expression provides insights into the mammalian microrna-processing machinery. *J Virol* 84(2):695–703
42. Godshalk SE, Bhaduri-McIntosh S, Slack FJ (2008) Epstein-barr virus-mediated dysregulation of human microrna expression. *Cell Cycle* 7(22):3595–3600
43. Cameron JE, Fewell C, Yin Q, McBride J, Wang X, Lin Z, Flemington EK (2008) Epstein-barr virus growth/latency iii program alters cellular microrna expression. *Virology* 382(2):257–266
44. Wang FZ, Weber F, Croce C, Liu CG, Liao X, Pellett PE (2008) Human cytomegalovirus infection alters the expression of cellular microrna species that affect its replication. *J Virol* 82(18):9065–9074
45. Okamura K, Phillips MD, Tyler DM, Duan H, Chou YT, Lai EC (2008) The regulatory activity of microrna* species has substantial influence on microrna and 3' utr evolution. *Nat Struct Mol Biol* 15(4):354–363
46. Okamura K, Liu N, Lai EC (2009) Distinct mechanisms for microrna strand selection by drosophila argonautes. *Mol Cell* 36(3):431–444
47. Tomari Y, Du T, Zamore PD (2007) Sorting of drosophila small silencing rnas. *Cell* 130(2):299–308
48. Lin EA, Kong L, Bai XH, Luan Y, Liu CJ (2009) Mir-199a, a bone morphogenic protein 2-responsive microrna, regulates chondrogenesis via direct targeting to smad1. *J Biol Chem* 284(17):11326–11335
49. Ro S, Park C, Young D, Sanders KM, Yan W (2007) Tissue-dependent paired expression of mirnas. *Nucleic Acids Res* 35(17):5944–5953

50. Bartel DP (2004) Micromnas: genomics, biogenesis, mechanism, and function. *Cell* 116(2):281–297
51. Bartel DP (2009) Micromnas: target recognition and regulatory functions. *Cell* 136(2):215–233
52. Umbach JL, Kramer MF, Jurak I, Karnowski HW, Coen DM, Cullen BR (2008) Micromnas expressed by herpes simplex virus 1 during latent infection regulate viral mmas. *Nature* 454(7205):780–783
53. Didiano D, Hobert O (2006) Perfect seed pairing is not a generally reliable predictor for mirna-target interactions. *Nat Struct Mol Biol* 13(9):849–851. doi:[nsmb1138](https://doi.org/10.1038/nsmb1138)[pii]10.1038/nsmb1138
54. Didiano D, Hobert O (2008) Molecular architecture of a mirna-regulated 3' utr. *RNA* 14(7):1297–1317. doi:[rna.1082708](https://doi.org/10.1261/rna.1082708)[pii]10.1261/rna.1082708
55. Vigorito E, Perks KL, Abreu-Goodger C, Bunting S, Xiang Z, Kohlhaas S, Das PP, Miska EA, Rodriguez A, Bradley A, Smith KG, Rada C, Enright AJ, Toellner KM, MacLennan IC, Turner M (2007) Micromna-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity* 27(6):847–859
56. Djeraba A, Bernardet N, Dambrine G, Quere P (2000) Nitric oxide inhibits marek's disease virus replication but is not the single decisive factor in interferon-gamma-mediated viral inhibition. *Virology* 277(1):58–65
57. He M, Xu Z, Ding T, Kuang DM, Zheng L (2009) Micromna-155 regulates inflammatory cytokine production in tumor-associated macrophages via targeting c/ebp β . *Cell Mol Immunol* 6(5):343–352
58. Wang X, Zhao Q, Matta R, Meng X, Liu X, Liu CG, Nelin LD, Liu Y (2009) Inducible nitric-oxide synthase expression is regulated by mitogen-activated protein kinase phosphatase-1. *J Biol Chem* 284(40):27123–27134
59. Schat KA, Nair V (2008) Marek's disease. In: SY M (ed) *Disease of poultry*, chapter 15, Neoplastic Diseases, pp 452–514
60. Murphy E, Vanicek J, Robins H, Shenk T, Levine AJ (2008) Suppression of immediate-early viral gene expression by herpesvirus-coded micromnas: implications for latency. *Proc Natl Acad Sci USA* 105(14):5453–5458

A p53-dependent promoter associated with polymorphic tandem repeats controls the expression of a viral transcript encoding clustered microRNAs

GRÉGOIRE STIK,¹ SYLVIE LAURENT,^{1,2} DAMIEN COUPEAU,¹ BAPTISTE COUTAUD,¹ GINETTE DAMBRINE,^{1,2} DENIS RASSCHAERT,¹ and BENOÎT MUYLKENS^{1,3}

¹Transcription, Lymphome Viro-Induit, University François Rabelais, UFR Sciences et Techniques, Parc de Grandmont, F-37200 Tours, France

²Department of Animal Health, INRA, F-37380 Nouzilly, France

³Embryology, Veterinary Department, Faculty of Sciences, University of Namur–FUNDP, B-5000 Belgium

ABSTRACT

The tumor suppressor protein p53 plays a role in cellular responses to cancer-initiating events by regulating progress through the cell cycle. Several recent studies have shown that p53 transactivates expression of the members of the proapoptotic microRNA-34 family, which are underexpressed in several cancers. We demonstrate here that the latency-associated cluster of microRNAs (miRNA) encoded by an oncogenic herpesvirus, *gallid herpesvirus 2* (GaHV-2), is a direct target of p53. Robust transcriptional activity was induced in three avian cell lines by a sequence mapping 600 base pairs (bp) upstream of the cluster of miRNAs. We found transcription start sites for the pri-miRNA transcript at the 3' end of this transcription-inducing sequence. The promoter has no consensus core promoter element, but is organized into a variable number of tandem repeats of 60-bp harboring p53-responsive elements (RE). The minimal functional construct consists of two tandem repeats. Mutagenesis to change the sequence of the p53 RE abolished transcriptional activity, whereas p53 induction enhanced mature miRNA expression. The identification of a viral miRNA promoter regulated by p53 is biologically significant, because all avirulent GaHV-2 strains described to date lack the corresponding regulatory sequence, whereas all virulent, very virulent, and hypervirulent strains possess at least two tandem repeats harboring the p53 RE.

Keywords: p53; promoter; microRNA; herpesvirus

INTRODUCTION

The transcription of eukaryotic genes results from a set of combinatorial events, from chromatin recognition to the positioning of the RNA polymerase at the appropriate initiation site. This process, which includes chromatin modification and the recruitment of cofactors and basal transcription components, leads to the assembly of an elongation-competent transcription complex. Some of the mechanisms underlying the activation of transcription have been elucidated, including those involving core promoter elements, such as the TATA box and initiator element (Sandelin et al. 2007). Several core elements of the TATA box containing human promoters have been identified; these elements include the downstream core

promoter element (Lee et al. 2005), the TFIIB recognition element (Deng and Roberts 2005) and the initiator element (INR). Mutagenesis of the INR sequence led to identification of the YYANWYY motif as a consensus sequence (Javahery et al. 1994; Smale and Kadonaga 2003). However, although the core promoter structure was initially thought to be invariant, it is now clear that there is extensive diversity. Thousands of different transcriptional start sites (TSSs) for mammalian class II promoters have been described, revealing the variability of potential TSSs (Frith et al. 2008). Transcription initiation may occur at various nucleotide positions within a core promoter region, but there is a preference for both a purine in position +1 and a pyrimidine in position –1 (Carninci et al. 2006; Frith et al. 2008). Otherwise, the proportion of promoters with clearly identifiable TATA boxes has tended to decrease as the number of promoters discovered increases (Suzuki et al. 2001; Sandelin et al. 2007; Yang et al. 2007). Thus, promoters lacking standard core promoter elements constitute a significant fraction of the eukaryotic genome. However, little is known about their mechanisms of

Reprint requests to: Denis Rasschaert, Transcription, Lymphome Viro-Induit, University François Rabelais, UFR Sciences et Techniques, Parc de Grandmont, F-37200 Tours, France; e-mail: denis.rasschaert@univ-tours.fr; fax: 33 024742 7774.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.2121210>.

transcriptional activation or the *cis*-activating DNA motifs involved.

MicroRNAs (miRNA) are a newly discovered class of small RNAs of ~23 nucleotides (nt) in length that have been shown to be key mediators of RNA silencing pathways (Ambros 2004; Bartel 2004; Ason et al. 2006). By fine-tuning gene expression through the repression of target mRNA translation, miRNAs control cellular pathways governing several biological functions and the fate of the cell: cell differentiation, cell proliferation, or apoptosis (Brennecke et al. 2003; Xu et al. 2003). Both the expression and function of miRNAs are tissue-specific and temporally regulated (Pasquinelli et al. 2005; Landgraf et al. 2007). Tumorigenesis is frequently associated with disruption of the differential and dynamic expression patterns of miRNAs (Tagawa et al. 2007; Chang et al. 2008). The spatiotemporal control of miRNA expression is thought to be principally transcriptional. However, the transcriptional regulation of miRNAs is still poorly understood, because few miRNA promoters have been characterized. Some transcription factors have been identified as positive (tumor suppressor protein 53, p53, NF- κ B) (Taganov et al. 2006; Chang et al. 2007) or negative (c-myc) (Chang et al. 2008) regulators of the expression of particular miRNAs. In one anti-tumoral pathway, p53 transactivates the pro-apoptotic miR-34a 30 kb downstream from the p53-responsive element (RE) (Chang et al. 2007; Raver-Shapira et al. 2007). Other roles of p53 in miRNA-dependent regulatory networks were discovered more recently, with the demonstration of p53-mediated transcriptional induction of miR-145 and miR-192 through the interaction of p53 with a potential p53 RE in their respective promoters (Song et al. 2008; Sachdeva et al. 2009). Otherwise, in a collection of 175 miRNAs for which transcription initiation regions have been described, 25 (19%) of the miRNA promoters contain a TATA element, 21% have a TFIIB recognition element, and 47% have a consensus initiator (Ozsolak et al. 2008). Thus, a significant proportion of miRNA genes lack standard core promoter elements and the transcriptional regulation of primary transcripts generating miRNAs (primiRNAs) is far from clear. Further and detailed investigation of the mechanisms governing miRNA expression is therefore required.

In this study, we investigated the transcriptional regulation of a viral microRNA cluster encoded within the latency-associated transcript (LAT) of *gallid herpesvirus 2* (GaHV-2), an alphaherpesvirus model of virus-induced lymphomagenesis (Burgess et al. 2004; Osterrieder et al. 2006). Thirteen viral miRNAs are expressed by oncogenic strains of GaHV-2, which cause T-cell lymphomas in chicken (Burnside et al. 2006; Burnside and Morgan 2007; Morgan et al. 2008; Xu et al. 2008). The viral miRNA coding sequences are clustered into two separate genomic loci: cluster 1 (mdv1-miR-M1, 2, 3, 4, 5, 9, 11, 12, 31) maps upstream of and downstream from the Meq gene; cluster 2 (mdv1-miR-M6, 7, 8, 10) maps near the 5' end of the LAT (Fig. 1). The LAT cluster encodes

four pre-miRNAs that are processed to give seven mature miRNAs (miR-M6-5p and 3p; miR-M7-5p and 3p; mdv1-miR-M8-5p and 3p; and miR-M10-5p) (Fig. 1). These miRNAs have been reported to be strongly expressed in MSB-1 cells derived from a GaHV-2 lymphoma and in primary MD lymphomas (Morgan et al. 2008; Yao et al. 2008). The LAT-miRNA coding sequences are located near the 5' end of a previously characterized large intron in LAT (Cantello et al. 1994, 1997). However, neither the promoter nor the TSS of LAT has been identified. Here, we identify and characterize a robust promoter controlling the expression of the LAT-miRNAs. This promoter was found embedded within a variable number of tandem repeats of the viral genome. No standard core promoter elements were found in this promoter, which was transactivated by p53 through a specific consensus RE.

RESULTS

Robust transcriptional activity is induced by the tandem repeat sequences upstream of the latency-associated transcript of GaHV-2 encoding viral miRNAs

We investigated the regulation of miRNA expression from the LAT, by studying putative promoter DNA sequences between the telomeric repeats and the locus encoding LAT-derived miRNAs (Fig. 1A,B). Genomatix software (www.genomatix.de/cgi) was used for in silico predictions of the *cis* response elements (RE) in this DNA sequence (Fig. 1C). The 5' region of this sequence amplified from RB-1B Bacmid DNA consists of three repeats of 60 base pairs (bp) containing only two types of RE (p53 and SpiB), with a number of potential REs predicted downstream (Oct-1, E2F, CAAT, and TATA boxes) (Fig. 1C). We defined the largest potential promoter, P1, between the first nucleotide downstream from the telomeric repeats of the IRS of GaHV-2 and the mdv1-miR-M8-5p sequence (extending from nucleotides 142486 to 143448 in the GaHV-2 reference genome, RB-1B). The promoter sequence was amplified by PCR from RB-1B Bacmid, and inserted into a promoter-less luciferase reporter plasmid (pGL3basic), to generate the P1 construct. The transcriptional activity associated with P1 was tested in three avian cell lines (LMH, DF-1, and 54-O) (Fig. 1D): P1 yielded robust reporter activity in all three cell lines (Fig. 1D), similar to that of the CMV promoter in the positive control (pcDNA luciferase; data not shown). We identified the sequence involved in the transcriptional activity of P1, by amplifying nested-set genomic fragments resulting from three deletions each at the 5' and 3' ends of P1 and testing them in luciferase reporter assays, as described above for P1 (Fig. 1C,D). Constructs Δ 1P1, Δ 2P1, and Δ 3P1 correspond to nested deletions from the 5' end of P1 deleting the 60-bp repeats. They had much lower levels of transcriptional activity than P1 (by factors of 20, 10 and 50 in LMH, DF1, and

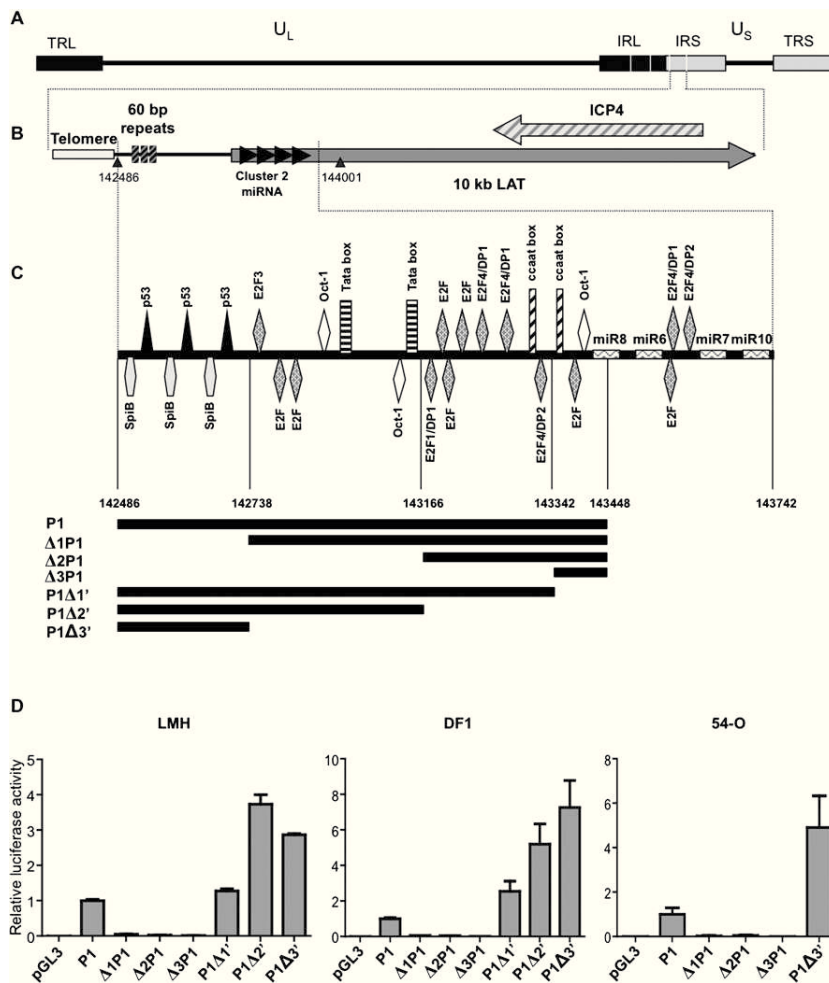


FIGURE 1. Localization of the latency-associated transcript promoter. (A) The organization of the GaHV-2 genome is similar to that of an E alpha herpes virus. UL and US indicate unique long and unique short sequences; TRL/TRS indicate terminal repeats, long and short, respectively; and IRL/IRS indicate internal repeats, long and short, respectively. (B) Localization of miRNA cluster 2 in the TRS and IRS. The gray arrow indicates a noncoding transcript and black arrows indicate cluster 2 miRNAs. (C) Schematic diagram of the various *cis* elements identified in the promoter by in silico analysis (Genomatix). Coordinates of the pre-mdv1-miRNAs are as follows: pre-mdv1-miR-M8 (143403-143482), pre-mdv1-miR-M6 (143524-143594), pre-mdv1-miR-M7 (143695-143772), and pre-mdv-miR-M10 (143815-143895). Potential promoter sequences are segmented and each potential promoter is represented in black under the schematic DNA sequence. (D) Relative luciferase activity for each promoter-luciferase construct, normalized with respect to the luciferase activity obtained with the P1 pGL3 vector is shown. These values reflect the transcriptional strength of the various promoters in LMH, DF1, and 54-O cells. Each value corresponds to the mean of three independent assays performed simultaneously. Error bars indicate the SEM for three replicates.

54-O cells, respectively). The P1Δ1', P1Δ2', and P1Δ3' constructs, corresponding to nested deletions from the 3' end of P1 and retaining all the 60-bp repeats had promoter activities similar to that of unmodified P1. Thus, the 60-bp repeats at the 5' end of P1 appear to be required for promoter activity, whereas the RE located downstream do not appear to be essential.

To confirm that the DNA promoter sequence was functional, we used rapid amplification of cDNA ends (RACE) to determine the 5' end of the LAT-miRNA transcript from the 54-O cell line, a GaHV-2 lymphoma cell line expressing high levels of LAT-encoded miRNAs (B. Muytjens, pers. comm.). All of the TSSs mapped to the last block of the 60-bp repeats or the proximal downstream region (Fig. 2A). We found that 16 of the 17 identified TSSs were spread over a 13-nt sequence that most of these TSSs mapped to the proximal downstream sequence (nucleotide +6, genomic coordinate 142734 in the RB-1B genome). Similarly, all of the TSSs corresponding to the RNA isolated from the LMH cell line transfected with the luciferase reporter vector P1Δ3' (chimeric LAT-Luc) mapped to this region (Fig. 2B). Thus, all the transcripts analyzed in two independent RACE experiments were found to have been initiated in the vicinity of the last block of 60-bp repeats, suggesting that the promoter of the LAT-encoded miRNAs is essentially located in the sequence corresponding to P1Δ3', from the telomeric sequences to the downstream proximal region of the last 60-bp repeat.

The promoter consisting of 60-bp repeats is responsible for the production of mature miRNAs from the LAT-miRNA-cluster2

The miRNA cluster is located within the LAT gene. We therefore investigated whether the promoter of this gene mediated the production of mature miRNAs. Two constructs, pB691 and pB695, encompassing the 5' end of the LAT mRNA defined above and the 4 miRNAs, with and without, respectively, the 60-bp repeats, were generated (Fig. 3A). The sequences corresponding to pB691 and pB695 were amplified from the bacmid

p-RB-1B, with primers binding 18 nt upstream of the first 60-bp repeat (M691) or in the terminal part of the last 60-bp repeat (M695), respectively, and 1719 nt downstream from miR10 (M656). These sequences were then inserted into the pBS-SK vector (Fig. 3A). The resulting constructs were then used to transfect LMH cells and the expression of MDV-miR-7-5p was estimated from Northern blots of extracted total

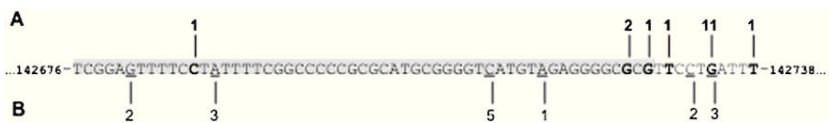


FIGURE 2. Identification of transcription start sites by 5' RACE-PCR. (A) Localization and frequencies of various TSSs identified by 5' RACE-PCR in 54-O cells, a cell line isolated from a lymphoma induced by GaHV-2. (B) Localization and frequencies of various TSSs identified by 5' RACE-PCR in LMH cells transfected with the P1Δ3' pGL3 vector (luciferase gene under the control of the P1Δ3' promoter). Transcription start sites are shown in bold for RNA from 54-O cells and are underlined for LMH cells transfected with the P1Δ3' pGL3 vector. Each TSS is labeled with the number of initiations found. The DNA sequence represents GaHV-2 genome, with the last 60-bp repeat highlighted in gray.

RNA with a radiolabeled probe specific for MDV-miR-7-5p, which is classically considered to be representative of miRNA cluster transcription (Yao et al. 2008). RNA from peripheral blood leukocytes (PBL) of GaHV-2-infected chickens over-expressing mdv1-miR-M7-5p during natural infection (B. Muylkens, pers. comm.) was used as a positive control, and RNA from nontransfected LMH cells was used as a negative control. A specific signal of the expected size for MDV-miR-7-5p was detected in LMH cells transfected with pB691, but not in LMH cells transfected with pB695 (Fig. 3B). Mature miRNA production therefore is strongly associated with the presence of the three 60-bp tandem repeats.

The core promoter of the LAT-encoded miRNAs consists of at least two 60-bp repeats

We found that the DNA sequence from the telomeres to the last 60-bp repeat upstream of the LAT locus displayed robust promoter activity, despite the absence of TATA and CAAT sequences. This region of tandem repeats in GaHV-2 isolates is hypervariable (Spatz and Silva 2007a): there are between 0 and 13 copies of the 60-bp direct repeat, depending on the viral strain. The P1Δ3' construct obtained from RB-1B Bacmide DNA contains three copies, with 63 nt of upstream and 10 nt of downstream sequences. We investigated the effects of both the sequences flanking the tandem repeats and the copy number of tandem repeats on promoter activity in the LMH and 54-O cell lines, by studying eight additional reporter constructs (Fig. 4). The reporter activities generated by these constructs were compared with that for the promoter construct P1Δ3' (arbitrarily set at 100%) (Fig. 4B,C). Removal of the sequence upstream of the tandem repeats did not significantly affect promoter

activity (construct P1Δ3'Tr5') (Fig. 4B,C). By contrast, removal of the short 3' sequence flanking the tandem repeats (construct P1Δ3'Tr3') reduced reporter activity to ~50% that of the P1Δ3' construct (Fig. 4B,C). This is consistent with the removal of one of the major TSSs identified within this sequence (Fig. 2). Deletion of both the 5' and 3' sequences surrounding the tandem repeats (construct [60 bp]₃) had no further effect on the activity of the promoter (Fig. 4B,C). Therefore, these surrounding se-

quences, although involved in positive regulation of the promoter, are not essential for core promoter function, because three direct repeats devoid of the surrounding sequences displayed >50% of the reporter activity of the highly active control construct.

We next investigated whether the number of copies of the repeats affected promoter activity (Fig. 4B,C). We used the MSB-1 cell line currently used as a reference, derived from a lymphoma induced by the virulent GaHV-2 strain BC-1, to obtain a construct with nine copies. The transcriptional activity of this construct was not markedly different from that of the reference construct P1Δ3', which has three copies of the tandem repeat. Constructs with two

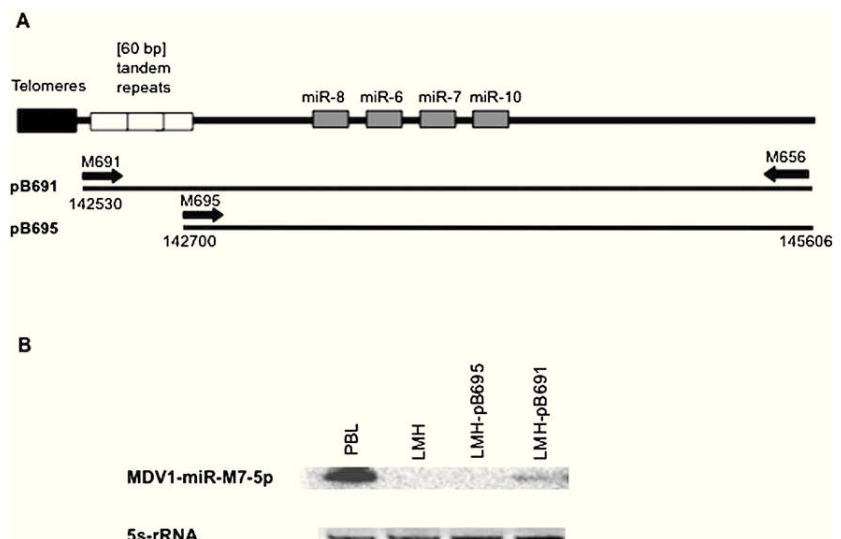


FIGURE 3. Northern blot analysis of the role of the core promoter in mature miRNA expression. (A) Schematic diagram of the LAT-miRNA promoter and the miRNA cluster. Each black line under the representation corresponds to a DNA fragment amplified from the bacmide p-RB-1B genome and inserted into the pBS-SK vector. (B) Northern blot analysis. Each lane was loaded with 15 µg of total RNA extracted with Trizol. Peripheral blood leukocytes were recovered 21 d after infection, from highly susceptible B13 chickens infected intramuscularly with RB-1B (1000 PFU of RB-1B). RNA was extracted from infected PBL, untransfected LMH cells and LMH cells transfected with pB691 or pB695. A ³²P-5' end-labeled DNA oligonucleotide probe complementary to the miRNA was used to detect the miR-7-specific signal. Ribosomal 5s RNA (ethidium bromide staining, EB) levels were used to check for equal loading.

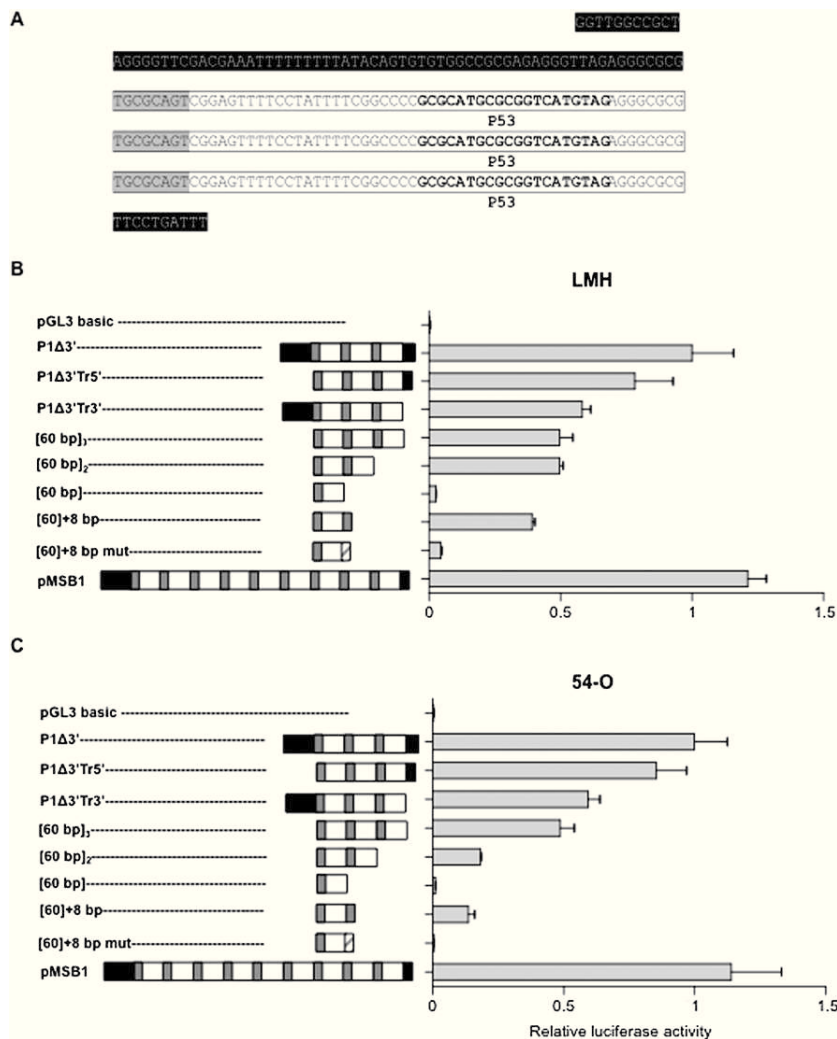


FIGURE 4. Determination of the core promoter of the LAT-miRNA coding sequence. (A) DNA sequence of the P1Δ3' promoter. White boxes correspond to the 60-bp repeats and sequences shaded in black correspond to nonrepeated sequences. Sequences shaded in gray correspond to the first eight nucleotides of a 60-bp block; p53 responsive elements are indicated in bold. (B,C) Relative luciferase activity obtained with various promoter constructs: [60 bp]₃, [60 bp]₂, or [60 bp]₁ carrying of three, two, or one 60-bp repeats, respectively, and [60 bp] + 8 bp and [60 bp] + 8 bp mut, contain one block of 60 bp with the first eight bases of a second repeat or eight mutated bases (polyG), respectively. The luciferase activities obtained with these constructs were normalized with respect to that of the P1Δ3' luciferase construct. Tests were performed with LMH (B) and 54-O (C) cells. Each value corresponds to the mean of three independent assays performed simultaneously. Error bars indicate the SEM of three replicates. Schematic diagrams of the various promoters are shown to the left of each histogram: sequences shaded in black correspond to nonrepeated sequences; white boxes correspond to the 60-bp repeats; gray boxes correspond to the first eight base pairs of the 60-bp repeat; dashed boxes correspond to the mutation of the first eight base pairs to eight guanines.

and three copies of the 60-bp repeat had 50% of the activity of the reference P1Δ3' construct in the LMH cell line, and no significant differences were observed between these two constructs (Fig. 4B). By contrast, the construct harboring a single copy of the 60-bp repeat displayed significantly lower levels of activity in LMH cells (3% that of P1Δ3'; $p <$

0.0001) (Fig. 4B). The effect of the number of copies of the 60-bp repeat on reporter activity was even more pronounced in the 54-O cell line (Fig. 4C): The construct with two copies displayed only 18% of the activity of P1Δ3'. The transcriptional activity of the construct with one copy did not exceed the background activity of a promoter-less, negative control plasmid (Fig. 4C). The addition of 8 bp at the 3' end of a single 60-bp repeat (construct [60 bp] + 8 bp) restored activity to 80% that for two full-length 60-bp repeats, in both LMH and 54-O cells (Fig. 4B,C). Invalidation of the 8-bp sequence at the end of construct ([60 bp] + 8 bp) by mutation severely impaired its activity (by 90% and 95% in LMH and 54-O cells, respectively) (Fig. 4B,C).

These findings indicate that the tandem repeat region upstream of the TSS of LAT requires a minimum of two copies of the 60-bp repeat for strong promoter activity. Nevertheless, a single 60-bp sequence with an 8-bp sequence from a second copy of the 60-bp repeat at the 3' end constituted a minimal functional construct.

The transcriptional activity of the LAT-miRNA promoter is dependent on the p53 responsive element

The sequence of the region of tandem repeats contains DNA stretches with similarities to transcription factor response elements (RE) and to potential transcription initiator signals (Fig. 5): the 60-bp sequence contains one potential M22 transcription initiator site (TGCGCANK-TGCGCAGT), one possible RE specific for SpiB (TTCC in reverse orientation), one canonical INR transcription initiator site (YYANWYY-CTATTTT), and one highly conserved consensus p53-binding site (GCGCAT GCGCGGTCATGTAG). The consensus

p53-binding sequence is a 20-bp DNA sequence consisting of two repeats of the RRRCWWGYYY motif separated by 0–13 nt. We investigated the possible involvement of these REs in the transcriptional activity of the 60-bp repeats, by evaluating their function with the shortest functional construct. This construct included a single copy of

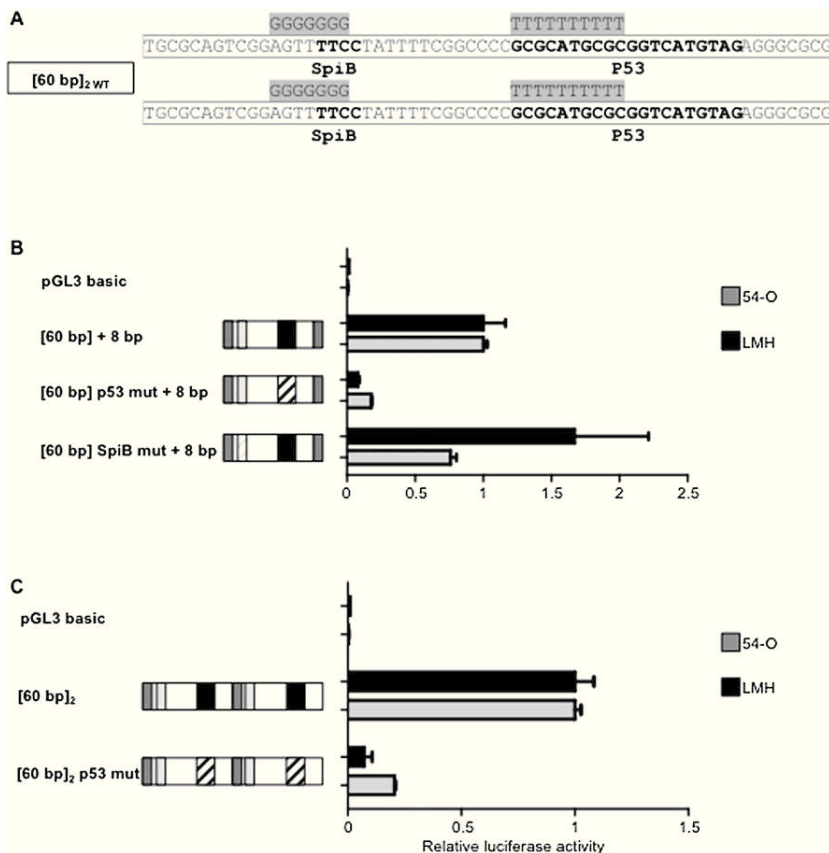


FIGURE 5. Identification of transcription factors involved in activation of the promoter. (A) Sequence of the core promoter [60 bp]₂, with various REs probably involved in transcriptional activity shown in bold. The various mutations of the [60 bp] repeats are shaded in gray. (B) Identification, by mutagenesis, of transcription factors involved in the function of the promoter. The [60 bp] + 8 bp promoter was used for mutagenesis. Each reported luciferase activity value is normalized with respect to that of the [60 bp] + 8 bp promoter construct. (C) Identification, by mutagenesis, of transcription factors involved in the function of the promoter. The [60 bp]₂ promoter was used for mutagenesis. Each reported luciferase activity value is normalized with respect to that of the [60 bp]₂ promoter construct. White boxes indicate 60-bp repeats containing the various RE sites and gray boxes correspond to the first eight base pairs of the 60-bp repeat. Black boxes correspond to the p53 RE, and light gray boxes correspond to the SpiB RE. Dashed boxes indicate mutated sites.

the 60-bp fragment followed by the first eight base pairs of a second copy ([60 bp] + 8 bp). We used site-directed mutagenesis to modify the two candidate REs and studied reporter activity in cells (Fig. 5B). The mutation of the p53 RE abolished 82% and 92% of the transcriptional activity in 54-O and LMH cells, respectively (Fig. 5B). By contrast, mutation of the SpiB RE did not significantly impair transcriptional activity in either cell line (Fig. 5B). We confirmed the effect of the p53 RE mutation in a construct containing two full-length copies of the 60-bp repeats, in which both p53 REs were invalidated (construct [60 bp p53mut]₂) (Fig. 5C). The transcriptional activity of construct [60 bp]₂ p53 mut was 7% that of the control in 54-O cells and 20% that for the control in LMH cells. These findings show that the p53 RE is directly involved in the transcriptional activity of the promoter.

p53 transactivates the promoter of the LAT-miRNA-cluster2

The minimal promoter construction, [60 bp]₂, has robust promoter activity, strongly associated with the presence of at least one p53 RE. We investigated the essential role of p53 in the transactivation of this core promoter, by three complementary approaches: (1) ChIP assays were performed from MSB-1 cells derived from a GaHV-2 lymphoma, (2) the [60 bp]₂ construct was used to transfect p53-deficient human osteosarcoma SAOS-2 cells complemented with human or chicken p53, and (3) the levels of three miRNAs from the cluster were measured in MSB-1 cells treated with adriamycin, a genotoxic drug that induces p53 accumulation, through phosphorylation of the Ser¹⁵ residue close to the mdm2-binding domain (Vassilev 2005).

For the first approach, we assessed the direct association between endogenous p53 and the LAT-miRNA promoter, by performing ChIP experiments with MSB-1 cells. Immunoprecipitated DNA was analyzed by PCR using specific primers, surrounding the p53 RE. We found that p53 was recruited to the LAT-miRNA promoter (Fig. 6). As expected, ladders resulting from primers hybridization to different consecutive repeats were observed after PCR amplification from immunoprecipitated DNA (Fig. 6).

For the second approach, two related sets of complementation experiments were carried out. SAOS-2 cells were cotransfected with the minimal construct [60 bp]₂ and an expression plasmid for chicken p53 (pcCHP53) or with the empty expression vector, pcDNA3.1(+), as control (Fig. 7A). The pcCHP53 construct was generated by inserting the chicken p53 cDNA downstream from the CMV promoter of the expression vector, pcDNA3.1(+). The ectopic expression of chicken p53 led to promoter activity of the minimal construct [60 bp]₂ twice as strong as that of the empty pcDNA3.1(+) vector control (Fig. 7A). The other complementation experiment used an expression plasmid for wild-type human p53 (pHP53SN) or for human p53 with a mutation in codon 143 (pHP53CX3) (Fig. 7B). Similar results were obtained: The promoter activity of the minimal construct [60 bp]₂ was 1.35 times higher when SAOS-2 cells were cotransfected with the expression plasmid for wild-type human p53 (pHP53SN) than when transfected with the

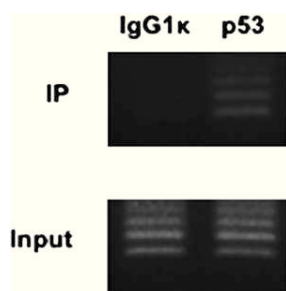


FIGURE 6. Direct association of p53 with the LAT-miRNA promoter. The PCR amplification is carried out using primers pair M645 and M735 (Table 1). Electrophoretic analyses of the PCR amplification products, using immunoprecipitated DNA (IP) or total input DNA (Input), are shown. Mouse IgG1 antibodies (*left* column) were used in ChIP assays as a negative control. The *right* column shows the PCR amplification products from DNA immunoprecipitated with p53 antibodies HP64. LAT-miRNA-promoter specific primers amplified products of various sizes from 154 bp to more with an increment of 60 bp corresponding to one 60-bp repeat.

construct encoding the mutated protein HP53CX3 (Fig. 7B). Residual promoter activity was observed in SAOS-2 cells transfected with the [60 bp]₂ construct alone. Other proteins of the p53 family have similar activities and the residual activity in SAOS-2 cells may be due to the effects of p73 or p63, which are not knocked out (Irwin and Kaelin 2001).

We evaluated the expression of three mature miRNAs of the LAT-miRNA-cluster2 (mdv1-miR-M6, mdv1-miR-M7, and mdv1-miR-M8) in MSB-1 cells treated with the genotoxic drug adriamycin, by quantitative RT-PCR, to confirm the p53-dependent activity of the promoter in natural conditions. The level of expression of mdv1-miR-M4, which belongs to the second cluster of MDV miRNAs, was determined as a negative control. Conversely, p21 mRNA levels were determined as a positive control of the transactivation effect dependent on the stabilization of p53 after adriamycin treatment. Western blot analysis of MSB-1 p53 expression after adriamycin treatment was carried out to prove the involvement of adriamycin in p53 stabilization. After 8, 12, and 24 h adriamycin treatment, the p53 expression level increases by a factor equal to 1.5, 2, and 3.8 respectively (Fig. 8B).

As expected, p21 mRNA transcription levels were 1.8 times higher in GaHV-2-infected MSB-1 cells treated with adriamycin for 8 h than in untreated MSB-1 cells. By contrast, the expression of mdv1-miR-M4, which does not belong to the LAT-miR-cluster2, was not affected by adriamycin treatment. Adriamycin increased the level of expression of the three miRNAs (mdv1-miR-M6, mdv1-miR-M7, and mdv1-miR-M8) of LAT-miRNA-cluster2 by factors of 1.8, 2.4, and 2, respectively (Fig. 8A). Same experiments performed in homolog cell line 54-O have shown similar results (data not shown). Collectively, these data provided strong evidence for the involvement of p53 in the production of mature LAT-miRNAs.

DISCUSSION

Herpesviruses have a large double-stranded DNA genome and have several conserved mechanisms for regulating gene expression, including complex alternative transcription and post-transcriptional regulation mechanisms triggered by miRNAs. They also produce noncoding RNAs (ncRNAs), which are involved in the establishment and maintenance of latency (Jones 2003). These latency-associated ncRNAs may be variants of long intranuclear noncoding RNAs antisense to various mRNAs described in mammals (Birney et al. 2007; Carninci 2008). In the avian oncogenic herpesvirus GaHV-2, LAT belongs to a family of RNAs spliced from a 10-kb ncRNA antisense to the mRNA for the major lytic transactivator protein ICP4 (Cantello et al. 1994, 1997). Recent studies have shown that LAT contains a cluster of miRNAs (Burnside et al. 2006; Burnside and Morgan 2007; Morgan et al. 2008; Xu et al. 2008). GaHV-2 lymphomagenesis occurs during the latency phase. The miRNA-LAT

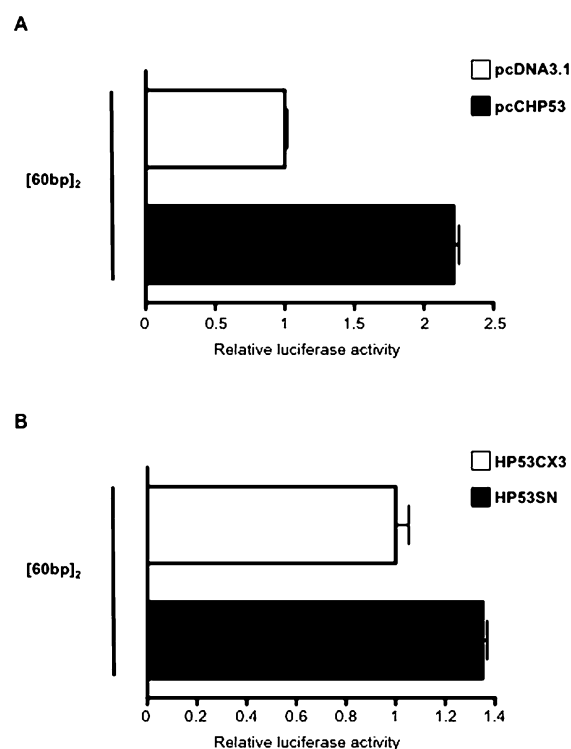


FIGURE 7. Tumor protein p53 transactivates the [60 bp]₂ minimal promoter. SAOS-2 cells were cotransfected with the minimal promoter [60 bp]₂ and an empty pcDNA3.1(+) expression plasmid and an expression plasmid for chicken p53 (pcCHP53) (A) or plasmids encoding wild-type human p53 (pHP53SN) and mutated human p53 (pHP53CX3) (B). Luciferase activities following cotransfection with pcCHP53 and pHP53SN were standardized with respect to those of pcDNA3.1 (A) and pHP53CX3 (B), respectively. PRL-TK (A) and PRL-CMV (B) vectors were used to normalize the transfection rate. Each value corresponds to the mean of at least three independent assays performed simultaneously. Error bars indicate the SEM for three replicates.

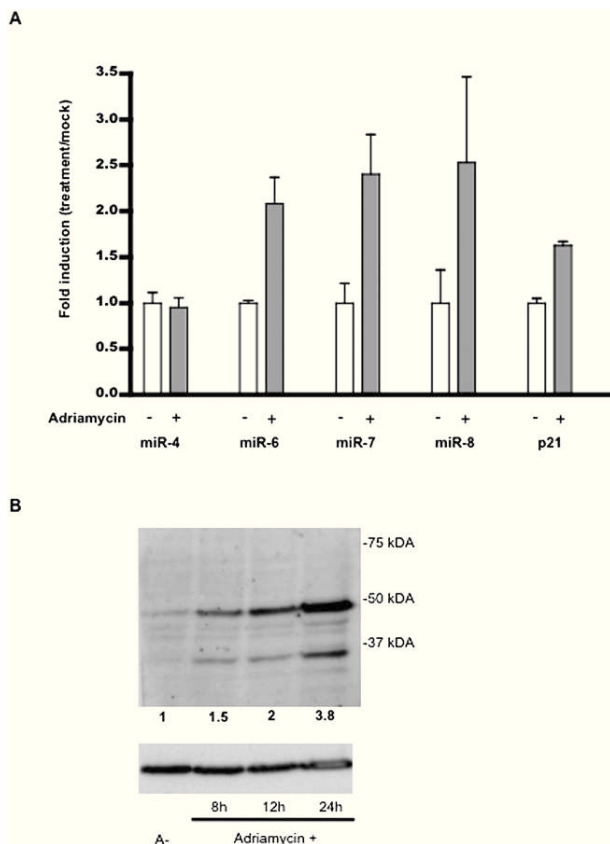


FIGURE 8. Induction of mature LAT-miRNA by adriamycin. (A) Adriamycin treatment induces mature LAT-miRNA expression. MSB-1 cells derived from a *Gallid herpes virus-2* (GaHV-2)-induced lymphoma were treated with adriamycin for 8 h. RNA was extracted with Trizol and analyzed by quantitative primer-extension PCR assay, with the Biorad analysis kit. Expression profiles for cluster 2 miRNAs (mdv1-miR-M7, -M8 and -M6) were recorded. The expressions of non-p53-specific mdv1-miR-M4 and p53-specific p21 were evaluated as negative and positive controls, respectively. Fold induction after adriamycin treatment is shown for each RNA. Each value corresponds to the mean of three independent assays performed simultaneously. Errors bars indicate the SEM of three replicates. (B) Western blot analysis of p53 and the GAPDH level after adriamycin treatment. MSB-1 cells were treated with adriamycin (0.2 μ g/mL) for 8, 12, and 24 h. Each blot was probed with mouse anti-p53 (HP64) and mouse anti-GAPDH primary antibodies, followed by rabbit anti-mouse horseradish-conjugated secondary antibody. GAPDH was used to check for equal protein loading. Relative signal intensities for p53 were calculated with respect to the GAPDH loading control, with Adobe Photoshop element 6 software. The values reported under the blots correspond to the respective level of p53 detected after adriamycin treatment, relative to the signal for p53 without adriamycin treatment.

gene is overexpressed in GaHV-2-transformed and latently infected cells, and it has therefore been suggested that the miRNA cluster is involved in latency and/or lymphomagenesis.

In this study, we successfully mapped the core promoter of this LAT-miRNA transcript 600 nt upstream of miRNA cluster 2 and showed that this promoter led to the transcrip-

tion of mature miRNAs (Fig. 3). The LAT-miRNA minimal promoter consists of at least two identical blocks of 60 bp. Each block contains one potential RE for p53 and one for SpiB. Only the p53 RE was found to be functional in the avian cell lines tested (Fig. 5). Its sequence (GCGCATGCG CGGTCATGTAG) is very similar to the consensus p53 RE sequence reported for humans and several other species. It is consistent with the 10-nt consensus RRRWWGYYY (half-site) tandem repeat sequences separated by 0–13 nt (el-Deiry et al. 1992). Systematic analysis of p53 REs and a recent *in vivo* study have shown that the most efficient promoters of the murine *Mdm2* gene have no spacer between the two half-sites of the p53 RE (Jordan et al. 2008). In light of these reports, the perfect palindrome for the 5' half-site and the lack of spacer bases between the two half-sites in the GaHV-2 miRNA-LAT promoter may explain its considerable transcriptional strength, almost as great as that of the strong CMV promoter.

A common feature identified in several p53 target genes is the presence of multiple p53-binding sites (el-Deiry et al. 1992; Resnick-Silverman et al. 1998; Hwang et al. 1999; Thornborrow and Manfredi 1999; Espinosa and Emerson 2001; Nakano and Vousden 2001). The distances between these p53 REs are variable. Synergies have been described between these multiple p53 REs (Thornborrow and Manfredi 1999; Jordan et al. 2008). However, the space between the p53 REs of the murine *mdm2* promoter cannot be significantly modified without a loss of function (Jordan et al. 2008). In the miRNA-LAT promoter described here, the p53 REs are 60 bp apart. The relative rates of transcription from the [60 bp] + 8 bp, [60 bp], and [60 bp]₂ constructs highlight the essential role of a transcription initiator sequence downstream from the p53 RE (Fig. 4), but we cannot exclude the possibility of synergy potentiating the p53 RE in neighboring 60-bp blocks. In lymphoid cells (54-O), derived from GaHV-2 lymphomas, the constructs harboring three full p53-binding sites displayed transcriptional activities twice as high as that of the construct possessing only two p53-binding sites. Furthermore, the promoter structure found to be functional in transcriptional analysis (Fig. 4) is biologically relevant, because genotyping data for a large panel of GaHV-2 strains showed that oncogenic strains possess at least two blocks of 60-bp harboring p53 REs (Spatz and Silva 2007b). However, the presence of more than three tandem repeats did not result in significantly greater transcriptional activity (Fig. 4).

The use of a reporter gene to analyze promoter activity showed that, in addition to the p53-binding site, one downstream functional INR played a significant role in the promoter activity of the LAT gene. The tandem repeats sequence contains a consensus initiator element (INR; CTATTTT) starting at nucleotide +19 of the 60-bp repeats. However, this INR appears to be a minor TSS, because only one (of 16) and three (of 17) transcripts started within this sequence in 5' RACE assays. By contrast, both 5'-end

analyses identified a major TSS of LAT in a sequence very similar to the INR consensus sequence YYANWYY, immediately downstream from the last 60-bp repeat (CTGATTT; position 142734 in the RB-1B genome). Several observations have implicated this INR-like sequence in promoter function. A construct based on three 60-bp repeats, but with the deletion of this downstream INR, displayed two-thirds the activity of an equivalent construct containing the INR motif in LMH cells and about one-third that of the same construct in 54-O cells. These findings were consistent with the numbers of TSSs found in these cell lines, as one-third of the TSSs in LMH cells and two-thirds of those in 54-O cells corresponded to this INR-like site. All the experiments indicated that a p53 RE was required upstream of any potential start site signal, for optimal transcription. Activity was retained if the INR-like sequence was replaced by the first eight nucleotides of the 5' end of the 60-bp repeat. These eight nucleotides correspond to the M22 site (TGCGCANK), which has been predicted to have a high prevalence in TATA-less core promoters (Birney et al. 2007).

Several examples of TATA-less promoters with various transcription start sites have been described (Carninci et al. 2006; Birney et al. 2007). Furthermore, studies on class II mammalian promoters have revealed that TSSs are closely clustered within the core promoter region (Carninci et al. 2006; Frith et al. 2008). We identified several TSSs in the last 58 bp of the promoter (Fig. 2). The TSSs of the chimeric LAT-Luc transcript are distributed throughout this 60-bp DNA motif, although most (94%) were found immediately downstream from the last 60-bp block near the INR-like CTGATTT in analyses of the LAT transcripts from the homologous system (Fig. 2); 64% of the TSSs were found at the classical position (third base) in this INR-like sequence (coordinate 142734 in the RB-1B genome). The LAT RNA is processed into spliced small RNAs by complex alternative splicing (Cantello et al. 1997), so the differences in the TSS distributions between the two systems may arise from this cotranscriptional event. Interestingly, detailed analysis of the LAT TSSs revealed that transcription started preferentially at sites with a purine in position +1 and a pyrimidine at position -1, as previously reported for the human and murine transcriptomes (Carninci et al. 2006). Indeed, 79% of TSSs (76% and 82% of the LAT-Luc and LAT transcripts, respectively) display a YR configuration at positions -1, +1 (Fig. 2). In addition, the composition of the surrounding sequence of most of the TSSs was consistent with recent observations based on exhaustive CAGE analyses (Frith et al. 2008).

Tumor suppressor protein p53 is a major transcription factor specifically activated in response to numerous stress signals, including those associated with cancers (Oren 2003; Harris and Levine 2005; Vousden and Prives 2005; Toledo and Wahl 2006). Alterations to this protein are among the most frequent causes of cancer. To date, three miRNA genes have been described as direct transcriptional targets of p53.

One, miR-34a, an miRNA from the miR-34 family involved in senescence, cell-cycle arrest, and apoptosis, maps 30 kb downstream from a p53-binding site. This miRNA displays p53-dependent up-regulation (Chang et al. 2007; Raver-Shapira et al. 2007). Several groups have reported that the miR-34 family, including miR-34a, miR-34b, and miR-34c, which target the Bcl-2, Notch, and HMGA2 mRNAs, respectively, are induced by DNA damage and oncogenic stress, in a p53-dependent manner (Tazawa et al. 2007; Ji et al. 2008). In 2008, miR-192 was also identified as a direct target of p53, suggesting that this miRNA is involved in cell-cycle regulation through the p53-miRNA pathways (Song et al. 2008). The recent identification of miR-145 as a direct p53 target has expanded the repertoire of p53-regulated genes. The induction of miR-145 by p53 leads to the down-regulation of c-Myc through miR-145-mediated silencing (Sachdeva et al. 2009). In our study, p53 seemed to trans-activate the miRNA-LAT promoter directly, through a specific p53 RE (Figs. 5–8). Previous studies revealed that LAT and miRNA cluster 2 in GaHV-2 are abundantly expressed during the latency and lymphomagenesis stages of viral infection (Burnside et al. 2006, 2008; Xu et al. 2008). Although p53 is considered to be a major tumor repressor, it has been shown that nonmutated p53 is present in GaHV-2-induced lymphomas (Gimeno et al. 2005). Further experiments are required to investigate the apparent contradiction between the production of native p53 and the development of lymphomas. The p53 cellular response to GaHV-2 infection may lead to overtranscription of the viral LAT-miRNA gene, potentially interfering with the pro-apoptotic pathways induced by p53.

Epidemiological observations support an association between the p53 target sequences driving the expression of the LAT locus and the oncogenic properties of GaHV-2. Several strains of GaHV-2 with different oncogenic capacities, correlated with their level of pathogenicity, have been isolated and partially sequenced. Recent published alignments of the sequences from a panel of 14 strains, including the promoter and part of the LAT gene, have shown that all the virulent, very virulent, and hypervirulent oncogenic GaHV-2 strains harbor two or more 60-bp tandem repeats (Spatz and Silva 2007a). By contrast, avirulent, mildly virulent, and vaccine strains have no copies of the tandem repeat. In addition, the minor effect on promoter activity of the short upstream sequence separating the telomeres from the 60-bp tandem repeats (Fig. 4) is reflected in vivo: this sequence has been lost from the genome of one of the very virulent strains (vv GaHV-2 strain 595).

In conclusion, we found sequences robustly initiating and inducing transcription of the pri-miRNA LAT transcript embedded within variable number tandem repeats harboring functional p53 binding sites in the GaHV-2 genome. The minimal functional construct consisted of two tandem repeats, consistent with the pathobiological characteristic of oncogenic and nononcogenic GaHV-2

strains. Transcriptional analysis indicated that a transcription initiator sequence downstream from a p53 binding site is required for the function of this TATA-less promoter. This p53-dependent promoter controlling viral miRNAs expression during viral latency and lymphomagenesis, allows the viruses to exploit the cellular p53 regulatory network.

MATERIALS AND METHODS

Bioinformatic analysis

Genomatix software (www.genomatix.de/cgi) was used for the in silico prediction of promoter sequences. This program identifies potential transcription factor response elements within a DNA sequence.

Cell culture, virus, and treatments

Five cell lines were used. The LMH line (ATCC no. CRL-217), established from chicken hepatic carcinoma epithelial cells, was cultured in DMEM:F12 medium (Cambrex) supplemented with 10% fetal bovine serum and 5% chicken serum. The avian fibroblast cell line DF1 (ATCC no. CRL-12203) was cultured in a DMEM medium similarly supplemented with sera. The MSB-1 cell line is derived from a spleen lymphoma induced by a virulent strain of MDV-1 (Akiyama et al. 1973). It was previously shown to be coinfecting with both MDV-1 (BC-1) and MDV-2 (HPRS24) (Hirai et al. 1990). The 54-O cell line was isolated in the TLVI laboratory (INRA, Nouzilly) from an ovary lymphoma induced by recombinant RB1-B virus reconstituted from an infectious genome cloned as a bacterial artificial chromosome (kindly provided by Venugopal Nair). These cell lines were cultured in RPMI-1640 medium (Cambrex) supplemented with 1 mM sodium pyruvate, 10% fetal bovine serum, and 5% chicken serum. Both latently infected MDV cell lines display comparable characteristics and express high levels of MDV microRNAs and the MSB-1 cell line is currently used as a reference. The p53-deficient osteosarcoma SAOS-2 cells (ATCC n°HTB-85) were cultured in a DMEM medium supplemented with 10% fetal bovine serum.

We treated cells with 0.2 µg/mL adriamycin for 8 h to induce p53 before the recovery of total RNA.

Plasmid constructs

The entire sequence of the putative promoter of the GaHV-2 LAT transcript was amplified from Bacmid DNA of the RB-1B strain (Petherbridge et al. 2003). The initial construct extended from nucleotide 142486 to nucleotide 143448 of the reference RB-1B genome (GenBank accession number EF523390), covering the region between the telomeric repeat and the LAT-miRNA cluster. Various primers were used to obtain different fragments of potential promoters (Fig 1; Table 1). PCR was performed as follows: 94°C for 4 min, followed by 25 cycles of denaturation (94°C 30 sec), annealing (55°C 30 sec), and extension (72°C 1 min), in a final volume of 50 µL containing 2.5 units of *Taq* DNA polymerase (NEB, New England Biolabs), 0.1 µM of each primer (Eurogentec), 0.2 mM of each deoxyribonucleotide, 10 mM Tris-HCl pH 9, 50 mM KCl, 0.5 mM MgCl₂, 0.1% TritonX-100, and 50 ng DNA.

A wild-type monomer of the 60-bp repeat was generated by PCR amplification of a 60-nt single-stranded oligonucleotide (Eurogentec). Multimers ([60 bp]₂ and [60 bp]₃) were constructed by annealing two partially complementary oligonucleotides of 60 nt. The resulting double-stranded fragments harbored overlapping ends that were used for concatemerization and ligation. The mutated promoters were generated in the same way, from 60-nt single-stranded oligonucleotides carrying mutations at the target sites. These sequences then amplified by PCR, with specific primers, in the conditions described above.

All the putative promoter sequences were amplified with primer pairs (Table 1) including *KpnI* and *HindIII* sites in their tails. The *KpnI/HindIII*-digested PCR products were inserted in the correct orientation between the *KpnI* and *HindIII* sites of the pGL3 basic vector (Promega).

The p53 expression vectors were provided by Thierry Soussi (UPMC/Karolinska Institutet). pHP53SN and pHP53CX3 encode the wild-type human p53 and a mutated form of human p53, respectively, under control of the CMV promoter. The chicken p53 cDNA was also provided by T. Soussi, flanked by *EcoRI* restriction sites in the pBS-SK plasmid. The p53 fragment was digested and inserted into the vector pcDNA3.1(+) linearized by digestion with *EcoRI*.

The two constructs, pB691 and pB695, encompassing the 5' end of the LAT mRNA and the four miRNAs, with and without, respectively, the 60-bp repeats, were amplified from the bacmid p-RB-1B, with primers binding 18 nt upstream of the first 60-bp repeat (M691) or in the terminal part of the last 60-bp repeat (M695), respectively, and 1719 nt downstream from miR10 (M656) (Table 1). Each primer used contains a *BamHI* restriction site in their tail. The *BamHI*-digested PCR products were inserted into the pBS-SK vector (Stratagene).

Plasmid DNA was purified with the Nucleobond Xtra midi kit (Macherey Nagel), and the inserts of all constructs were completely sequenced.

Luciferase assays

The pGL3basic (Promega) expression vector was used to test the strength of promoters inserted upstream of the *firefly* luciferase (F-Luc) gene. The pRL-TK (Promega) vector includes an expression cassette for the *Renilla* luciferase (R-Luc) under the control of the thymidine kinase promoter of human herpes virus 1 (HHV-1). Cotransfection with these two vectors (pGL3 carrying the test promoter and pRL-TK) made it possible to normalize measurements as a function of the rate of transfection. The pRL-CMV vector contains an expression cassette for the *Renilla* luciferase (R-Luc) under the control of the CMV promoter. The use of this vector also makes it possible to normalize the transfection rate.

Lipofectamine reagent (Invitrogen) was used for the transfection of adherent cells. Transfection assays were performed with 2×10^4 cells per well, in 96-well plates, with a mixture of 300 ng of firefly luciferase reporter construct and 5 ng of *Renilla* luciferase control vector.

We transfected 54-O lymphoblastoid cells with plasmids by electroporation. We washed 54-O cells twice with RPMI medium and adjusted the cell density to 8×10^6 cells/ 800 µL of serum-free medium. For each electroporation, 40 µg of the

TABLE 1. Primers used to construct different sequences of promoters and to perform RACE-PCR

Orientation-number-name	Sequence (5'-3')
F-M645-P1	TCAGGGTACCGCTAGGGGTTTCGACGAAATTT
F-M647-Δ1P1	TCAGGGTACCGTAGAGGGCGGTTCTGATT
F-M648-Δ2P1	CGACGGTACCTGGTGTGGCGCCAAAAATGC
F-M649-Δ3P1	CGGAGGTACCGGATCCCCGATCCACATATCC
R-M651	TAGCAAGCTTACGAGAACGATAGATCGAAACC
R-M735-P1Δ3'	GCATAAGCTTAAATCAGGAACGCGCCCTCTAC
R-M736-P1Δ2'	GCATAAGCTTGCATTTTTTGGCGCCAACACCA
R-M737-P1Δ1'	GCATAAGCTTGGATATGTGGATCGGGAATCG
F-M762-[60 bp]	GCCGGGTACCTGCGCAGTCGGAGTTTTCTATTTT
R-M763-[60 bp] + 8 bp	GCGGAAGCTTACTGCGCACGCGCCCTCTACATGAC
F-M787-[60 bp] SpiBmut + 8 bp	GCCGGGTACCTGCGCAGTCGGGGGGGGCTATTTT
R-M778-[60 bp]	GCGGAAGCTTTCGCGCCCTCTACATGACCGCGC
R-M779-[60 bp] + 8 bp mut	GCGGAAGCTTTTTTTTTTCGCGCCCTCTACATGACC
R-M780-[60 bp] ₂ p53 mut	GCGGAAGCTTCGCCCTCTACATGACCAAAAAA
F-M760-[60 bp]	TGCGCAGTCGGAGTTTTCTATTTTCGGCCCCGCGCATGCGCGGTCATGTAGAGGGCGCG
R-M761-[60 bp] + 8 bp	ACTGCGCACGCGCCCTCTACATGACCGCGCATGCGCGGGGCGGAAAAATAGGAAAACTCCG
F-M764-[60 bp] p53 mut	TGCGCAGTCGGAGTTTTCTATTTTCGGCCCCTTTTTTTTTGGTCATGTAGAGGGCGCG
R-M765-[60 bp]p53mut + 8 bp	ACTGCGCACGCGCCCTCTACATGACCAAAAAAAGGGGCGGAAAAATAGGAAAACTCCG
F-M781-[60 bp] SpiB mut	TGCGCAGTCGGGGGGGGCTATTTTCGGCCCCGCGCATGCGCGGTCATGTAGAGGGCGCG
R-M782-[60 bp] SpiBmut + 8 bp	ACTGCGCACGCGCCCTCTACATGACCGCGCATGCGCGGGGCGGAAAAATAGCCCCCCCCCG
GeneRacer-RNA oligo	CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUAGAAA
F-GeneRacer 5' oligo	CGACTGGAGCACGAGGACACTGA
F-GeneRacer 5' nested oligo	GGACACTGACATGGACTGAAGGAGTA
R-M803-GeneRacer LUC	CCAGGAACCAGGGCGTATCTCTT
R-M804-geneRacerLUCnested	CCAGCGTTCCATCTTCCAGCGGATA
R-M714-GeneRacer LAT	CGGATGCTGGAGCTGCCGCCAAACTTG
R-M810-GeneRacerLATnested	CTGCGGTAATCGGAATCGAGTTC
F-M695-MinigeneP-	GTCGACGGATCCCCGCGCATGCGCGGTCATGTAGAGGGCGCG
F-M691-MinigeneP+	GGTACC GGATCCATACAGTGTGTGGCCGCGAGAGGGTT
R-M656-Minigene	ATAACCGGATCCCGGGATGTTGTGCGGCCATTAACTCTCA
qmiR-M6-5P	CATGATCAGCTGGGCCAAGAACGAGAACACT
qmiR-M7-5p	CATGATCAGCTGGGCCAAGATCGGGATCTCC
qmiR-M8-5p	CATGATCAGCTGGGCCAAGATCGAAACCAAC
qmiR-M4-5p	CATGATCAGCTGGGCCAAGAACGAAGGGTTT
UP-qPCRmiR-Universal primer	CATGATCAGCTGGGCCAAGA
LNA-miR-M6-5p	CZGTTPTTCCGTAGT
LNA-miR-M7-5p	TPTTAZCTCGGGGA
LNA-miR-M8-5p	TETPTTCTGTGGT
LNA-miR-M4-5p	TZAATPCTGTATCG
p21Sense	CGAGCAGATCCAGAACGACTT
P21AntiSense	CGATGAGCCCCACCGACACCA
UP-ASmiR-M6-5P	CATGATCAGCTGGGCCAAGAACGAGAACACTACGGAACAACA
UP-ASmiR-M7-5p	CATGATCAGCTGGGCCAAGATCGGGATCTCCCCGAGATAACA
UP-ASmiR-M8-5p	CATGATCAGCTGGGCCAAGATCGAAACCAACCAGAACAAATA
UP-ASmiR-M4-5p	CATGATCAGCTGGGCCAAGAACGAAGGGTTCCGATACAGCATTAA
AS-MiR-M7-5P	TCGGGATCTCCCCGAGATAACA

Restriction endonuclease sites used for insertion into the vector are underlined. Mutated sites are shown in bold. LNA-modifications: E = A-LNA, L = C-LNA, P = G-LNA, and Z = T-LNA.

firefly luciferase reporter construct and 600 ng of *Renilla* luciferase control vector were mixed with cells. The cells were then electroporated in a 4 mm cuvette subjected to 100 V/mm for 5 ms. The contents were then dispensed into 2.5 mL of complete medium.

Luciferase assays were performed 24 h after transfection, with the dual luciferase reporter assay system (Promega) and a Centro LB960 luminometer (Berthold Technologies). Student's *t*-test was used for statistical analysis.

We transfected SAOS-2 cells in 24-well plates seeded with 1×10^5 cells per well, using Lipofectamine reagent (Invitrogen). Each well was transfected with a mixture of 1 μ L Lipofectamine, 400 ng promoter construct, 100 ng inductive expression plasmid (p53chicken pcCHP53, human wild-type HP53SN or human mutated HP53CX3), and 10 ng of pRL-TK or pRL-CMV (for cotransfection with chicken and human p53 constructs, respectively). For SAOS-2 cells, we carried out the dual luciferase reporter assay 12 h after transfection.

5' RACE analysis

5' Rapid amplification of cDNA ends (5' RACE) was carried out on total RNA isolated from two cell cultures. The first RNA sample was isolated from 54-O cells and the second was isolated from LMH cells transfected with pGL3-P1Δ3' (containing the P1Δ3' promoter). Total RNA was extracted with the Trizol extraction kit. cDNAs were obtained with the GeneRacer kit (Invitrogen). This technique is based on RNA ligase-mediated and oligo-capping RACE methods, and results in the selective ligation of an RNA oligonucleotide to the 5' ends of decapped RNA. The resulting cDNA was amplified by PCR, using forward primers annealing to the ligated RNA oligonucleotide and reverse primers annealing to the target gene (Table 1).

The 5' end of the ligated cDNA obtained from 54-O cells was determined with reverse primers binding to a published exon of *LAT*. The 5' end of the ligated cDNA obtained from transfected LMH cells was determined with reverse primers binding to the firefly luciferase ORF. PCR products of various sizes were obtained and inserted into the cloning vector PCR4-TOPO TA. We sequenced 20 and 42 positive clones for LMH and 54-O cells, respectively, to determine the 5' end in each case.

Northern blot analysis

RNA was extracted in Trizol and 15 µg of total RNA per lane was resolved by electrophoresis in a 15% acrylamide/bisacrylamide (19:1) gel containing 7 M urea, in Tris-borate-EDTA (TBE) buffer. Gels were briefly stained with ethidium bromide, and RNA was transferred, by electroblotting, onto a Nylon Plus membrane (Amersham) for 1.5 h at 350 mA in 0.50X TBE. The membrane was cross-linked by exposure to UV (Stratalinker, Stratagene). The membrane was subjected to prehybridization for 1 h and was then hybridized overnight in Perfect Hyb TM Plus hybridization buffer (Sigma) at 50°C, with a 20 nM ³²P-5' end-labeled DNA oligonucleotide probe complementary to the miRNA. It was then washed in low-stringency wash buffer. Blots were analyzed by phosphorimaging with a FLA7000 scanner from Fuji.

ChIP assays

ChIP assays were carried out as previously described (Shkreli et al. 2007). Briefly, chromatin from 10⁷ MSB-1 cells was cross-linked, washed, resuspended in lysis buffer and sonicated 12 times 6 sec with an 18J-pulse (VibraCell Bioblock Scientific). A fraction of the total chromatin was taken as the total input DNA control. After preclearing with protein G agarose/salmon sperm DNA beads, protein/DNA complexes were immunoprecipitated overnight at 4°C with anti-p53 (70 µL of hybridomas supernatant of HP64 antibody kindly provided by T. Soussi) or with mouse immunoglobulin G1κ (6 µg of IgG1κ, Sigma Aldrich) as a negative control. Immunoprecipitated complexes were collected with Protein G beads. Immunoprecipitation products were washed with low salinity buffer, high-salinity buffer, lithium chloride washing buffer, and finally with TE (Tris-EDTA) buffer. Immunocomplexes were extracted and cross-linking was reversed by incubation at 65°C overnight. DNA fragments were then purified using Nucleospin Extract II (Macherey-Nagel) and eluted in 30 µL of ultrapure water. PCR were performed in 50 µL volume with Promega Taq DNA polymerase and with primers pair M645/

M735 (Table 1). PCR was initiated with 3 min at 94°C followed by 35 cycles of 94°C for 30 sec, 55°C for 30 sec, 72°C for 1 min, and finally 7 min at 72°C.

Western blot analysis

MSB-1 cell lines were treated with 0.2 µg/mL of adriamycin during 8, 12, and 24 h, respectively. Five hundred thousand cells were treated with lysis buffer (15 mM Tris, 15 mM NaCl, 20 mM EDTA, 1% SDS, 5% beta-mercaptoethanol, 20% glycerol, bromophenol blue, [pH 6.8]). Proteins were separated by electrophoresis in 10% polyacrylamide gels containing SDS and passively blotted onto nitrocellulose membranes. The membranes were blocked by incubation with 3% skim milk powder in PBST (PBS, 0.5% Tween 20) and probed with monoclonal mouse anti-p53 (HP64 antibody kindly provided by T. Soussi) and monoclonal mouse anti-GAPDH (Millipore) primary antibodies, followed by polyclonal rabbit anti-mouse horseradish peroxidase-conjugated secondary antibodies. The immunoreactions of interest were detected with the Femto Super Signal Western detection kit (Pierce Chemical Co.). GAPDH was used as a control to demonstrate equal protein loading. Adobe Photoshop element 6 software was used to calculate the signal intensities for p53 level after different timing induction relative to the GAPDH loading control.

miRNA quantification by q-RT-PCR

Mdv1-miR-M4-5p, -M6-5p, -M7-5p, and -M8-5p RNA levels were quantified as described by Raymond et al. (2005). Briefly, total RNA was isolated by guanidinium thiocyanate-phenol-chloroform extraction (TRIzol, Invitrogen). After total RNA quantification by Nanodrop methods, the cDNAs were synthesized in primer extension assays containing 1 µg of total RNA (Table 1). Following reverse transcription, triplicate measurements were made on 2 µL of cDNA in a final reaction volume of 20 µL, by qPCR, in a 96-well optical PCR plate, with an iCycler qPCR device (Biorad). SYBR green PCR mixture, containing 10 µL iQ SYBR green PCR master mix (Biorad), 5 µL water, 1 µL of 10 mM universal primer, 1 µL of 10 mM LNA-R primer, and 3 µL of sample. The primer sequences are reported in Table 1. qPCR was carried out in the conditions recommended by the manufacturer, and dissociation curves were generated post-run for the analysis of amplicon species. The temperature at the dissociation peak is reported in Table 1. For the detection and quantification of chicken p21 mRNA, we used the SYBR green method with p21 sense and antisense oligomers as primers (Table 1), as described elsewhere (Sato et al. 2006).

ACKNOWLEDGMENTS

We thank Dr. Nair Venugopal (IAH, Compton) for providing us with the Bacmid p-RB1B and Sascha Trapp for the 54-O cell line. We also thank Dr. Thierry Soussi (UPMC; Karolinska), who kindly provided us with the antibody anti-p53 (HP64), the expression vectors for wild-type human p53 (pHP53SN) and mutated human p53 (pHP53CX3), and the cDNA for chicken p53. This work was supported by the "Ligue Nationale contre le cancer, Comité du Cher, Comité de l'Indre" and the "Agence Nationale de la Recherche" (ANR-07-MIME-012-01). B.M. is a postdoctoral researcher supported by the *Fonds National de la Recherche Scientifique* (FNRS).

Received February 10, 2010; accepted August 13, 2010.

REFERENCES

- Akiyama Y, Kato S, Iwa N. 1973. Continuous cell culture from lymphoma of Marek's disease. *Biken J* **16**: 177–179.
- Ambros V. 2004. The functions of animal microRNAs. *Nature* **431**: 350–355.
- Ason B, Darnell DK, Wittbrodt B, Berezikov E, Kloosterman WP, Wittbrodt J, Antin PB, Plasterk RH. 2006. Differences in vertebrate microRNA expression. *Proc Natl Acad Sci* **103**: 14385–14389.
- Bartel DP. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297.
- Birney E, Stamatoiyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**: 799–816.
- Brennecke J, Hipfner DR, Stark A, Russell RB, Cohen SM. 2003. bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell* **113**: 25–36.
- Burgess SC, Young JR, Baaten BJ, Hunt L, Ross LN, Parcells MS, Kumar PM, Tregaskes CA, Lee LF, Davison TF. 2004. Marek's disease is a natural model for lymphomas overexpressing Hodgkin's disease antigen (CD30). *Proc Natl Acad Sci* **101**: 13879–13884.
- Burnside J, Morgan RW. 2007. Genomics and Marek's disease virus. *Cytogenet Genome Res* **117**: 376–387.
- Burnside J, Bernberg E, Anderson A, Lu C, Meyers BC, Green PJ, Jain N, Isaacs G, Morgan RW. 2006. Marek's disease virus encodes microRNAs that map to meq and the latency-associated transcript. *J Virol* **80**: 8778–8786.
- Burnside J, Ouyang M, Anderson A, Bernberg E, Lu C, Meyers BC, Green PJ, Markis M, Isaacs G, Huang E, et al. 2008. Deep sequencing of chicken microRNAs. *BMC Genomics* **9**: 185. doi: 10.1186/1471-2164-9-185.
- Cantello JL, Anderson AS, Morgan RW. 1994. Identification of latency-associated transcripts that map antisense to the ICP4 homolog gene of Marek's disease virus. *J Virol* **68**: 6280–6290.
- Cantello JL, Parcells MS, Anderson AS, Morgan RW. 1997. Marek's disease virus latency-associated transcripts belong to a family of spliced RNAs that are antisense to the ICP4 homolog gene. *J Virol* **71**: 1353–1361.
- Carninci P. 2008. Non-coding RNA transcription: Turning on neighbours. *Nat Cell Biol* **10**: 1023–1024.
- Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engstrom PG, Frith MC, et al. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* **38**: 626–635.
- Chang TC, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KH, Feldmann G, Yamakuchi M, Ferlito M, Lowenstein CJ, et al. 2007. Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol Cell* **26**: 745–752.
- Chang TC, Yu D, Lee YS, Wentzel EA, Arking DE, West KM, Dang CV, Thomas-Tikhonenko A, Mendell JT. 2008. Widespread microRNA repression by Myc contributes to tumorigenesis. *Nat Genet* **40**: 43–50.
- Deng W, Roberts SG. 2005. A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes Dev* **19**: 2418–2423.
- el-Deiry WS, Kern SE, Pietenpol JA, Kinzler KW, Vogelstein B. 1992. Definition of a consensus binding site for p53. *Nat Genet* **1**: 45–49.
- Espinosa JM, Emerson BM. 2001. Transcriptional regulation by p53 through intrinsic DNA/chromatin binding and site-directed co-factor recruitment. *Mol Cell* **8**: 57–69.
- Frith MC, Valen E, Krogh A, Hayashizaki Y, Carninci P, Sandelin A. 2008. A code for transcription initiation in mammalian genomes. *Genome Res* **18**: 1–12.
- Gimeno IM, Witter RL, Fadly AM, Silva RF. 2005. Novel criteria for the diagnosis of Marek's disease virus-induced lymphomas. *Avian Pathol* **34**: 332–340.
- Harris SL, Levine AJ. 2005. The p53 pathway: Positive and negative feedback loops. *Oncogene* **24**: 2899–2908.
- Hirai K, Yamada M, Arao Y, Kato S, Nii S. 1990. Replicating Marek's disease virus (MDV) serotype 2 DNA with inserted MDV serotype 1 DNA sequences in a Marek's disease lymphoblastoid cell line MSB1-41C. *Arch Virol* **114**: 153–165.
- Hwang BJ, Ford JM, Hanawalt PC, Chu G. 1999. Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair. *Proc Natl Acad Sci* **96**: 424–428.
- Irwin MS, Kaelin WG. 2001. p53 family update: p73 and p63 develop their own identities. *Cell Growth Differ* **12**: 337–349.
- Javahery R, Khachi A, Lo K, Zenzie-Gregory B, Smale ST. 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol* **14**: 116–127.
- Ji Q, Hao X, Meng Y, Zhang M, Desano J, Fan D, Xu L. 2008. Restoration of tumor suppressor miR-34 inhibits human p53-mutant gastric cancer tumorspheres. *BMC Cancer* **8**: 266. doi: 10.1186/1471-2407-8-266.
- Jones C. 2003. Herpes simplex virus type 1 and bovine herpesvirus 1 latency. *Clin Microbiol Rev* **16**: 79–95.
- Jordan JJ, Menendez D, Inga A, Nourredine M, Bell D, Resnick MA. 2008. Noncanonical DNA motifs as transactivation targets by wild type and mutant p53. *PLoS Genet* **4**: e1000104. doi: 10.1371/journal.pgen.1000104.
- Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, et al. 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**: 1401–1414.
- Lee DH, Gershenzon N, Gupta M, Ioshikhes IP, Reinberg D, Lewis BA. 2005. Functional characterization of core promoter elements: The downstream core element is recognized by TAF1. *Mol Cell Biol* **25**: 9674–9686.
- Morgan R, Anderson A, Bernberg E, Kamboj S, Huang E, Lagasse G, Isaacs G, Parcells M, Meyers BC, Green PJ, et al. 2008. Sequence conservation and differential expression of Marek's disease virus microRNAs. *J Virol* **82**: 12213–12220.
- Nakano K, Vousden KH. 2001. PUMA, a novel proapoptotic gene, is induced by p53. *Mol Cell* **7**: 683–694.
- Oren M. 2003. Decision making by p53: Life, death and cancer. *Cell Death Differ* **10**: 431–442.
- Osterrieder N, Kamil JP, Schumacher D, Tischer BK, Trapp S. 2006. Marek's disease virus: From miasma to model. *Nat Rev Microbiol* **4**: 283–294.
- Ozsolak F, Poling LL, Wang Z, Liu H, Liu XS, Roeder RG, Zhang X, Song JS, Fisher DE. 2008. Chromatin structure analyses identify miRNA promoters. *Genes Dev* **22**: 3172–3183.
- Pasquinelli AE, Hunter S, Bracht J. 2005. MicroRNAs: A developing story. *Curr Opin Genet Dev* **15**: 200–205.
- Petherbridge L, Howes K, Baigent SJ, Sacco MA, Evans S, Osterrieder N, Nair V. 2003. Replication-competent bacterial artificial chromosomes of Marek's disease virus: Novel tools for generation of molecularly defined herpesvirus vaccines. *J Virol* **77**: 8712–8718.
- Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, Moskovits N, Bentwich Z, Oren M. 2007. Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. *Mol Cell* **26**: 731–743.
- Raymond CK, Roberts BS, Garrett-Engle P, Lim LP, Johnson JM. 2005. Simple, quantitative primer-extension PCR assay for direct monitoring of microRNAs and short-interfering RNAs. *RNA* **11**: 1737–1744.
- Resnick-Silverman L, St Clair S, Maurer M, Zhao K, Manfredi JJ. 1998. Identification of a novel class of genomic DNA-binding sites

- suggests a mechanism for selectivity in target gene activation by the tumor suppressor protein p53. *Genes Dev* **12**: 2102–2107.
- Sachdeva M, Zhu S, Wu F, Wu H, Walia V, Kumar S, Elble R, Watabe K, Mo YY. 2009. p53 represses c-Myc through induction of the tumor suppressor miR-145. *Proc Natl Acad Sci U S A* **106**: 3207–3212.
- Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA. 2007. Mammalian RNA polymerase II core promoters: Insights from genome-wide studies. *Nat Rev Genet* **8**: 424–436.
- Sato F, Kurokawa M, Yamauchi N, Hattori MA. 2006. Gene silencing of myostatin in differentiation of chicken embryonic myoblasts by small interfering RNA. *Am J Physiol Cell Physiol* **291**: C538–C545.
- Shkreli M, Dambrine G, Soubieux D, Kut E, Rasschaert D. 2007. Involvement of the oncoprotein c-Myc in viral telomerase RNA gene regulation during Marek's disease virus-induced lymphomagenesis. *J Virol* **81**: 4848–4857.
- Smale ST, Kadonaga JT. 2003. The RNA polymerase II core promoter. *Annu Rev Biochem* **72**: 449–479.
- Song B, Wang Y, Kudo K, Gavin EJ, Xi Y, Ju J. 2008. miR-192 regulates dihydrofolate reductase and cellular proliferation through the p53-microRNA circuit. *Clin Cancer Res* **14**: 8080–8086.
- Spatz SJ, Silva RF. 2007a. Polymorphisms in the repeat long regions of oncogenic and attenuated pathotypes of Marek's disease virus 1. *Virus Genes* **35**: 41–53.
- Spatz SJ, Silva RF. 2007b. Sequence determination of variable regions within the genomes of gallid herpesvirus-2 pathotypes. *Arch Virol* **152**: 1665–1678.
- Suzuki Y, Tsunoda T, Sese J, Taira H, Mizushima-Sugano J, Hata H, Ota T, Isogai T, Tanaka T, Nakamura Y, et al. 2001. Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res* **11**: 677–684.
- Taganov KD, Boldin MP, Chang KJ, Baltimore D. 2006. NF- κ B-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc Natl Acad Sci* **103**: 12481–12486.
- Tagawa H, Karube K, Tsuzuki S, Ohshima K, Seto M. 2007. Synergistic action of the microRNA-17 polycistron and Myc in aggressive cancer development. *Cancer Sci* **98**: 1482–1490.
- Tazawa H, Tsuchiya N, Izumiya M, Nakagama H. 2007. Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. *Proc Natl Acad Sci* **104**: 15472–15477.
- Thornborrow EC, Manfredi JJ. 1999. One mechanism for cell type-specific regulation of the bax promoter by the tumor suppressor p53 is dictated by the p53 response element. *J Biol Chem* **274**: 33747–33756.
- Toledo F, Wahl GM. 2006. Regulating the p53 pathway: In vitro hypotheses, in vivo veritas. *Nat Rev Cancer* **6**: 909–923.
- Vassilev LT. 2005. p53 Activation by small molecules: Application in oncology. *J Med Chem* **48**: 4491–4499.
- Vousden KH, Prives C. 2005. p53 And prognosis: new insights and further complexity. *Cell* **120**: 7–10.
- Xu P, Vernoooy SY, Guo M, Hay BA. 2003. The *Drosophila* microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. *Curr Biol* **13**: 790–795.
- Xu H, Yao Y, Zhao Y, Smith LP, Baigent SJ, Nair V. 2008. Analysis of the expression profiles of Marek's disease virus-encoded microRNAs by real-time quantitative PCR. *J Virol Methods* **149**: 201–208.
- Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E. 2007. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* **389**: 52–65.
- Yao Y, Zhao Y, Xu H, Smith LP, Lawrie CH, Watson M, Nair V. 2008. MicroRNA profile of Marek's disease virus-transformed T-cell line MSB-1: Predominance of virus-encoded microRNAs. *J Virol* **82**: 4007–4015.

RESEARCH ARTICLE

Open Access

The chicken miR-150 targets the avian orthologue of the functional zebrafish MYB 3'UTR target site

Audrey Guillon-Munos^{1,2†}, Ginette Dambrine^{1,3†}, Nicolas Richerieux^{1,4}, Damien Coupeau¹, Benoît Muylkens^{1,5}, Denis Rasschaert^{1*}

Abstract

Background: The *c-myb* proto-oncogene is the founding member of a family of transcription factors involved principally in haematopoiesis, in diverse organisms, from zebrafish to mammals. Its deregulation has been implicated in human leukaemogenesis and other cancers. The expression of *c-myb* is tightly regulated by post-transcriptional mechanisms involving microRNAs. MicroRNAs are small, highly conserved non-coding RNAs that inhibit translation and decrease mRNA stability by binding to regulatory motifs mostly located in the 3'UTR of target mRNAs conserved throughout evolution. MYB is an evolutionarily conserved miR-150 target experimentally validated in mice, humans and zebrafish. However, the functional miR-150 sites of humans and mice are orthologous, whereas that of zebrafish is different.

Results: We identified the avian mature miRNA-150-5P, *Gallus gallus* gga-miR-150 from chicken leukocyte small-RNA libraries and showed that, as expected, the gga-miR-150 sequence was highly conserved, including the seed region sequence present in the other miR-150 sequences listed in miRBase. Reporter assays showed that gga-miR-150 acted on the avian MYB 3'UTR and identified the avian MYB target site involved in gga-miR-150 binding. A comparative *in silico* analysis of the miR-150 target sites of MYB 3'UTRs from different species led to the identification of a single set of putative target sites in amphibians and zebrafish, whereas two sets of putative target sites were identified in chicken and mammals. However, only the target site present in the chicken MYB 3'UTR that was identical to that in zebrafish was functional, despite the additional presence of mammalian target sites in chicken. This specific miR-150 site usage was not cell-type specific and persisted when the chicken *c-myb* 3'UTR was used in the cell system to identify mammalian target sites, showing that this miR-150 target site usage was intrinsic to the chicken *c-myb* 3'UTR.

Conclusion: Our study of the avian MYB/gga-miR-150 interaction shows a conservation of miR-150 target site functionality between chicken and zebrafish that does not extend to mammals.

Background

c-myb was originally identified as the chicken cell homologue of the *v-myb* oncogenes found in two strains of avian leukaemia virus [1,2]. These avian *v-myb* oncogenes induce myeloid and erythroid forms of leukaemia in chickens and the activation of the *c-myb* promoter by the insertion of avian and murine retroviruses has also

been implicated in diverse forms of leukaemia [3,4]. A role for MYB in human leukaemogenesis was initially suspected following the demonstration of MYB overproduction in cells from patients with leukaemia. This role has recently been confirmed by the detection of duplications and translocations affecting the *c-myb* locus, particularly in acute and chronic myeloid leukaemia and in acute T-cell lymphoblastic leukaemia [5,6]. MYB deregulation is also associated with colorectal cancers [7,8], carcinomas [9] and breast cancers expressing oestrogen receptor- α [10], in which MYB has been implicated in prolactin-induced signalling pathways [11].

* Correspondence: denis.rasschaert@univ-tours.fr

† Contributed equally

¹Université François Rabelais, Equipe Transcription, Lymphome Viro-induit, UFR Sciences et Techniques, Parc de Grandmont, F-37200 Tours, France
Full list of author information is available at the end of the article

In normal cells, MYB has been shown to be essential for haematopoietic lineage specification, T- and B-lymphocyte differentiation, colonic mucosal crypt regeneration and brain neurogenesis, on the basis of the abnormal phenotypes observed in mouse *myb* mutants [12,13]. In zebrafish, MYB has also been shown to be essential for haematopoiesis [14,15] and the silencing of *c-myb* in zebrafish embryos also leads to abnormal phenotypes, with effects on eye tissue formation in particular [16].

c-myb is the founding member of a family of genes encoding transcription factors with a DNA-binding domain consisting of three regions: R1, R2 and R3 [12]. Vertebrate genomes contain two other closely related genes from this family [17]: *MYBL1* (also known as *A-myb*), which is expressed in a restricted panel of tissues, and *Mybl2* (also known as *B-myb*), which is ubiquitously expressed. The products of these genes regulate the expression of genes involved in the control of cell proliferation and differentiation [18-20]. In plants, the R2R3 MYBs constitute a large family of transcription factors involved in the regulation of plant-specific developmental and physiological processes [21].

MYB generally acts as a transcriptional activator, binding to the MYB binding site on DNA [22] and recruiting the CBP/p300 coactivator proteins [23,24], thereby increasing the level of transcription of MYB targets. The oncogenic properties of MYB result from the overexpression or inappropriate activation of *c-myb* [13].

The level of *c-myb* expression is critical and is regulated by a number of mechanisms. At the transcriptional level, the *c-myb* promoter may be transactivated by a large number of proteins, including MYB itself, and attenuation sequences identified in the first intron of *myb* also regulate Myb elongation [13,25]. No particular pattern of cell specificity has been clearly demonstrated for transcriptional regulation processes, whereas tight control over the levels of mature MYB mRNA is restricted to tissue compartments with a high turnover. MYB mRNA and MYB protein have a very short half-life [13]. They are present in large amounts in haematopoietic progenitor cells but are absent from terminally differentiated cells, suggesting that rapid changes occur [26]. An absolute requirement for the fine-tuning of *c-myb* expression was recently highlighted by the demonstration of a compromising effect of decreases in *c-myb* gene activity on murine haematopoietic stem cells, leading to a myeloproliferative disorder involving stem cells with novel characteristics [27].

The post-transcriptional mechanisms regulating MYB levels involve microRNAs (miRNAs). miRNAs constitute a class of highly conserved small (21-24 nucleotides) non-coding RNAs found in plants and animals [28]. In animals, mature miRNAs generated by processing from the stem loop pre-miRNA are incorporated into the

silencing complex (for a review see [28]), which mediates post-transcriptional repression by binding to mRNA molecules, causing a decrease in the rate of translation or stability of the target mRNA [28-32]. Metazoan miRNA target recognition is based on Watson-Crick pairing of the 5' region of the miRNA, mostly via nucleotides 2 to 8 – known as the miRNA seed – to sites generally located within the 3'-untranslated region (UTR) of the target mRNA. Several hundreds of miRNAs with highly conserved sequences have been identified in diverse vertebrate species, from fish to humans [28]. Moreover, the lengths and sequences of the 3' UTRs from mammals and birds are sufficiently similar for alignments to be generated [33]. Prediction algorithms (reviewed in [29]) based on base-pairing between the miRNA seed and mRNA sequences have identified potential miRNA target sites in thousands of human genes [34,35]. A comparative analysis of several mammalian genomes suggested that miRNA target sites were conserved regulatory motifs in mammals [33]. A pioneering comparative analysis of mammalian and fish miRNA targets yielded 240 orthologous miRNA target genes conserved between these two groups [34].

It has been predicted that miRNAs may regulate the production of proteins from as many as 10% to 30% of the genes present in the human genome [35]. There is evidence to suggest that miRNA function is critical for normal cellular development and homeostasis. The production of miR-150 in mature B and T cells has been shown to block early B-cell development, and its ectopic production in haematopoietic stem cells and progenitor cells has been shown to result in significantly lower than normal numbers of mature B cells [36]. Moreover, miR-150 controls B-cell differentiation by targeting murine MYB in a dose-dependent manner [37]. By targeting MYB, miR-150 also drives the differentiation of murine megakaryocyte-erythrocyte progenitors into megakaryocytes, indicating a key role for miR-150 in controlling lineage commitment [38]. Thrombopoietin also induces megakaryopoiesis by downregulating MYB expression through the effects of miR-150 [39]. Overall, in haematopoietic lineages and the B-cell subsets of tonsil tissues, miR-150 and MYB display opposite patterns of expression [40]. Opposite patterns of miR-150 and MYB expression are also observed in immortalised cell lines, in which no miR-150 is detected [41].

Like most vertebrate mRNAs [35], MYB mRNAs are conserved targets of microRNAs. Functional miR-150 target sites have been validated in human, murine and zebrafish MYB 3'UTRs [16,37] and orthologous functional target sites of miR-150 have been identified in mammals, but are not conserved in zebrafish.

We sought to increase knowledge about miR-150/MYB interactions during the course of evolution, by

investigating the function of avian miR-150. Several attempts at extensive or deep sequencing [42-47] have generated a list of more than a hundred chicken miRNAs in miRbase, but only a few chicken targets have been validated [48-50]. The chicken, an amniote that has evolved separately from mammals for about 310 million years, has provided unique data on vertebrate evolution [51]. The evolutionary distance between chicken and mammals may eventually facilitate studies of the target site adaptation of mammalian miRNAs. We used a reporter assay to identify the functional targets of miR-150 in the avian MYB 3'UTR and performed a comparative *in silico* analysis of predicted and experimentally validated functional target sequences in the vertebrate MYB 3'UTR (frog, fishes, chicken, mammals). Our findings show that the number of putative miR-150 target sites has increased during evolution, with chicken and mammals displaying an additional set of target sites, but with only one of the sites in chicken being functional, that identical to the functional site in zebrafish. The avian miR-150 bound to the same target sites in chicken cells and in the human cell line used to identify mammalian miR-150 target sites, demonstrating the specificity of this pattern of binding to the avian MYB 3'UTR.

Results and Discussion

Cloning of mature avian gga-miR-150

We cloned avian mature miRNA-150-5P (*Gallus gallus*, gga-miR-150) from small-RNA libraries derived from spleen or peripheral blood leukocytes before and at advanced stages of Marek's disease virus (MDV)-induced lymphomagenesis in chickens (69 reads). By contrast, gga-miR-150 was not detected in libraries derived from chicken immortalised lymphoid cell lines [52,53]; [our unpublished data]. The expression pattern of the avian gga-miR-150 is similar to that of the human hsa-miR-150, which is expressed in normal human haematopoietic cell lineages but not in immortalised cell lines [16,41]. The sequences of miRNAs seem to have been conserved during evolution, because the mature gga-miR-150 cloned from our libraries and the recently released sequence [54]<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL6541> contain the seed region (nt2 to nt8) sequence (CUCCCAA) identified in the other twelve miR-150 sequences listed in miRBase [55,56]<http://microrna.sanger.ac.uk> (Fig. 1). All mature miR-150 sequences begin with a uracil (U) residue and only four positions have been found to harbour changes (Fig. 1). The mature gga-miR-150 sequence differed from its human and murine counterparts by one residue, and from the zebrafish sequence by two residues (Fig. 1).

Chicken gga-miR-150 targets avian MYB

We investigated the targeting of avian MYB by gga-miR-150 in luciferase reporter assays on cell cultures. The sequence of the precursor pri-miRNA gga-miR-150 is not present in the released sequence of the *Gallus gallus* genome [57]<http://www.ncbi.nlm.nih.gov/projects/genome/guide/chicken/> and we were unable to identify pri-miRNA gga-miR-150 on the basis of synteny with human and zebrafish. We therefore constructed a synthetic avian pre-miR-150, based on a comparison of various mammalian pre-miR-150 sequences and stem loop structures (Additional file 1). The synthetic avian pre-miR-150 amplicon obtained by PCR with appropriate primers (Additional file 2) was introduced into a pcDNA3 vector (pmiR-150). We used a miR-150 mutant plasmid (pmiR-150mut) as a control, with the deletion of one C residue within the miR seed sequence (Fig. 1), to prevent pairing between gga-miR-150 and putative targets in the avian MYB 3'UTR. We assessed the functionality of gga-miR-150, by establishing stable lymphoid chicken MDV cell lines ectopically producing gga-miR-150 or mutated gga-miR-150. The chicken MDV cell lines MSB-1 [58] and PA9 [59] were transfected with pmiR-150, pmiR-150mut or pcDNA vectors. MDV cells ectopically producing gga-miR-150 (MDVT-150) or mutated gga-miR-150 (MDVT-150mut) and control pcDNA-transfected cells (MDVT) were then isolated by selection on neomycin. Northern blots showed the ectopic production of a mature gga-miR-150 of the same size by MDVT-150 cells and spleen cells, whereas a slightly smaller mature mutated gga-miR-150 was detected in MDVT-150mut cells (Fig. 2A). These findings are consistent with the deletion of one nucleotide from the mature mutated gga-miR-150, confirming the successful transcription and maturation of synthetic avian pre-miR-150 from vectors with pol II promoters [60].

We investigated the targeting of avian MYB by gga-miR-150, by inserting the 3'UTR fragment of the chicken MYB cDNA into a pRL-TK vector downstream from the *Renilla* luciferase reporter gene. The various MDVT cells were cotransfected with the MYB WT pRL-TK vector and the internal control pcDNAMluc vector encoding firefly luciferase, by electroporation, and dual luciferase assays were performed. We normalised luciferase activity, taking the value for MDVT-150mut cells as 100%, as mutated gga-miR-150 did not repress activity in MDVT-150mut cells. The activity of the MYB WT reporter was specifically repressed by gga-miR-150 in PA9 and MSB-1T-150 cells (by up to about 80%) (Fig. 2B). We also transiently cotransfected avian DF1 fibroblasts with pmiR-150, pmiR-150mut or pcDNA plasmids and the MYB WT pRL-TK vector. Reporter assays in

hsa-miR-150	UCUCCCAACCCUUGUACCAGUG
ptr-miR-150	UCUCCCAACCCUUGUACCAGUG
mml-miR-150	UCUCCCAACCCUUGUACCAGUG
mmu-miR-150	UCUCCCAACCCUUGUACCAGUG
rno-miR-150	UCUCCCAACCCUUGUACCAGUG
cfa-miR-150	UCUCCCAACCCUUGUACCAGUG
eca-miR-150	UCUCCCAACCCUUGUACCAGUG
bta-miR-150	UCUCCCAACCCUUGUACCAGUG <u>U</u>
mdo-miR-150	UCUCCCAACCCUUGUACCAG <u>AGU</u>
oan-miR-150	UCUCCCAACCCUUGUACCAG <u>AG</u>
gga-miR-150	UCUCCCAACCC <u>C</u> UGUACCAGUG
dre-miR-150	UCUCCCA <u>AU</u> CCUUGUACCAGUG
xtr-miR-150	UCUCCCAACCCUUGUACCAG <u>AG</u>
gga-miR-150mut	UCU-CCAACCC <u>C</u> UGUACCAGUG

Figure 1 Comparison of the sequence of avian gga-miR-150 with all miR-150 sequences from the miRBase. The human sequence was taken as the reference sequence. The seed sequence is shown in red and changes are shown in green, underlined. Sequences were predicted on the basis of sequence similarity or experimentally validated. Hsa, *Homo sapiens*; ptr, *Pan troglodytes*; mml, *Macaca mulatta*; mmu, *Mus musculus*; rno, *Rattus norvegicus*; cfa, *Canis familiaris*; eca, *Equus caballus*; bta, *Bos taurus*; mdo, *Monodelphis domestica*; oan, *Ornithorhynchus anatinus*; gga, *Gallus gallus*; dre, *Danio rerio*; xtr, *Xenopus tropicalis*. The sequence of the mutated gga-miR-150 (gga-miR-150mut) that we constructed is shown.

these cells confirmed that pmiR-150 efficiently reduced MYB WT reporter activity to about 60% (Fig. 3). Thus, in chickens, as in humans [16], the repression of MYB WT reporter activity by miR-150 does not seem to be cell type-specific and gga-miR-150, like its orthologues in humans, mice and zebrafish [16,37], specifically targets the 3'UTR of MYB.

gga-miR-150 uses a target site orthologous to that of zebrafish

We identified four putative miR-150 target sites in the chicken MYB 3'UTR (Fig. 3A), all matching the criteria for miRNA target recognition and target site sequence context within the 3'UTR [35,61-63]. Sites s1 and s3, located 85-106 nt and 830-850 nt downstream from the stop codon, respectively, were 7-mer-A1 sites containing the seed sequence with an additional A residue in target position 1 (Fig. 3A). Such sites have been shown to be functional in various systems [63]. Sites s2 and s4, located 123-142 nt and 918-938 nt downstream from the stop codon, respectively, were 9-mer-A1 and 8-mer-A1 sites (Fig. 3A). These sites are probably at least as effective as the 8-mer sites shown to be functional in various systems, from worms to mammals [29]. Additional pairing to the middle and the 3' end of gga-miR-

150 was also observed, extending from one individual match to seven consecutive matches, for sites s4 and s2, respectively (Fig. 3A). We generated MYB mutant pRL-TK vectors, each containing only one of the four mutated target sites (mt1 to mt4), and one mutant containing all four mutated target sites (mt1234). The various MYB mutant pRL-TK vectors were used to transfect either avian DF1 fibroblasts [64] or human 293FT cells, as described above. Avian DF-1 fibroblasts were used to assess gga-miR-150 site usage in the homologous system. We used human 293FT cells, originating from the HEK-293T cell line, as a reference heterologous system, because the HEK-293T cell line has been used to identify miR-150 target sites in mice, humans and zebrafish [16,37]. We then carried out reporter assays. The same reporter activity patterns were observed in both systems, demonstrating that the observed pattern of target site usage resulted from intrinsic characteristics of chicken *c-myb* 3'UTR. We found that pmiR-150 did not repress the activity of the MYB mt1234 reporter or the MYB mt2 construct whereas the activities of MYB mt1, mt3 and mt4 were significantly repressed (by about 40 to 60%) (Fig. 3B). Thus, gga-miR-150 specifically targets the 3'UTR of avian MYB, acting principally through target site s2 in

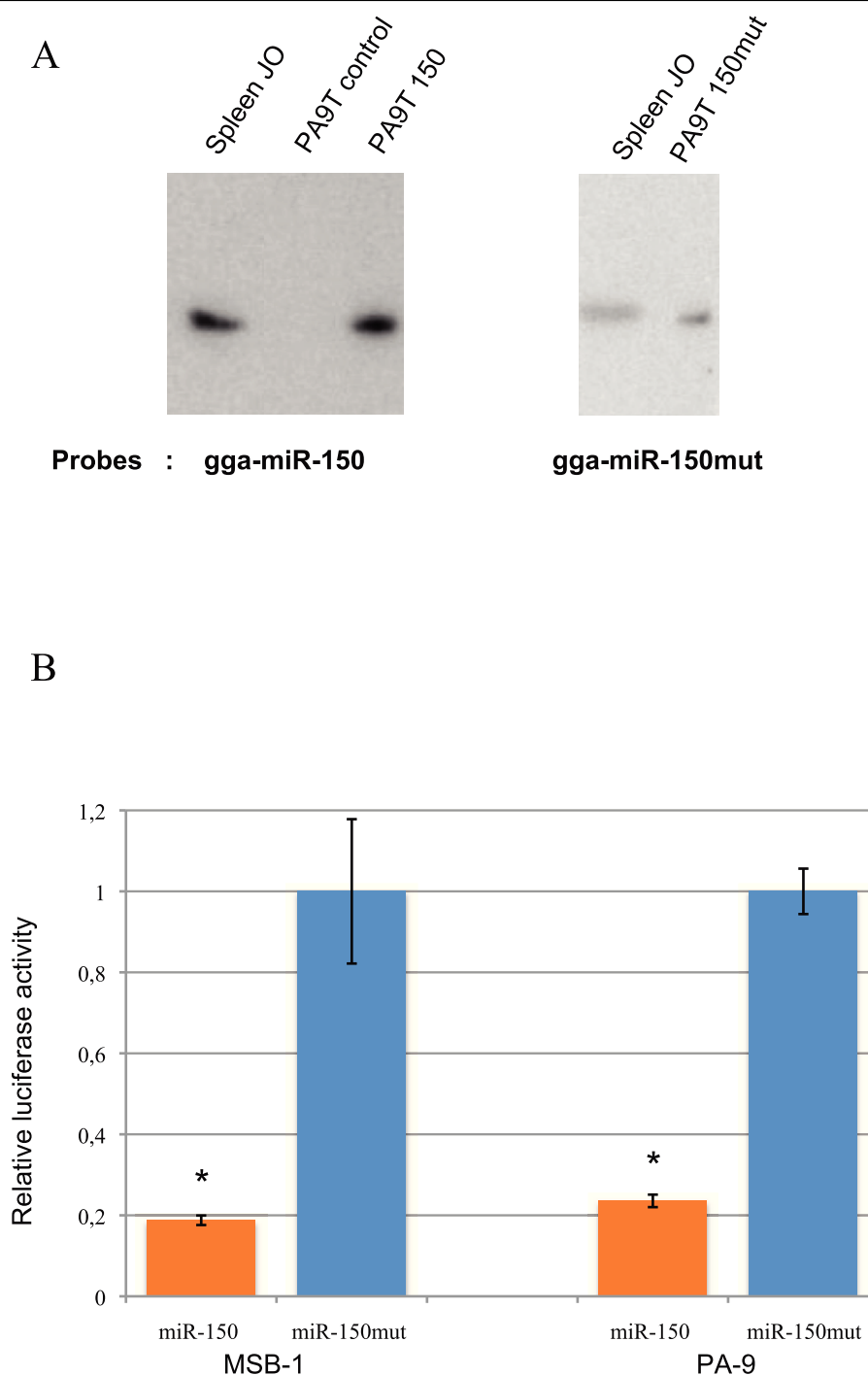
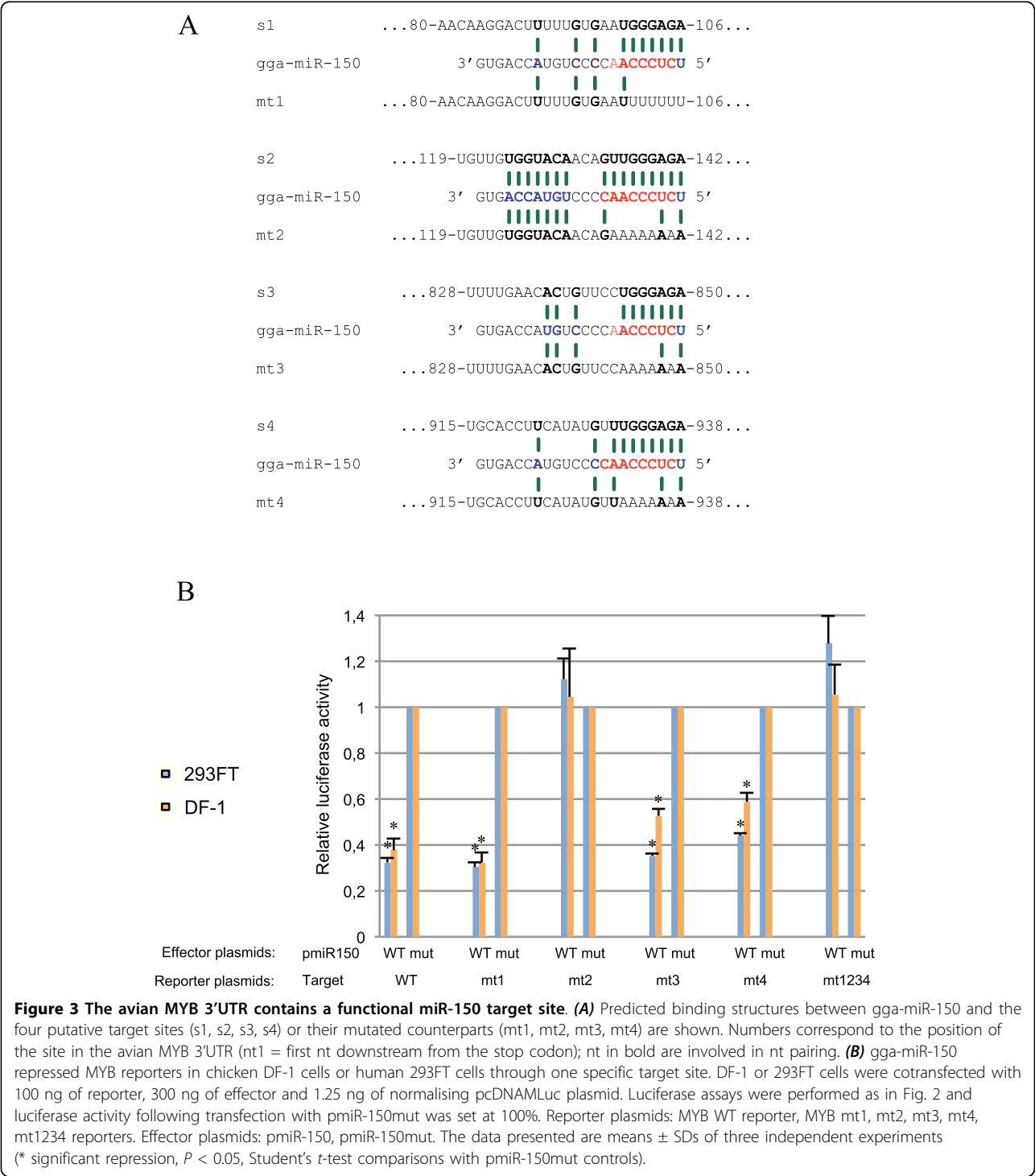


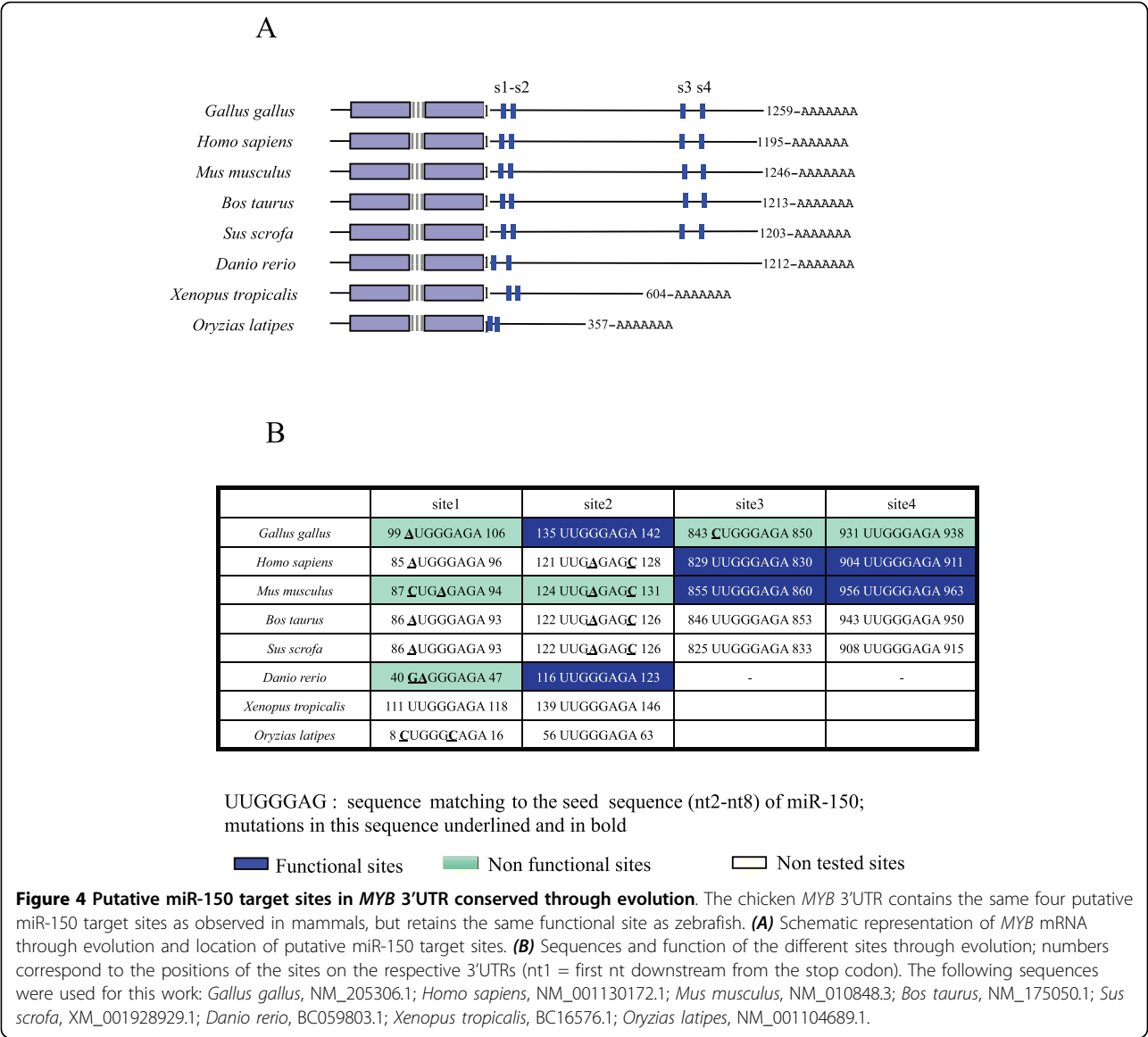
Figure 2 gga-miR-150 ectopically produced in stable lymphoid chicken PA9 and MSB-1 cells represses the avian MYB reporter. (A) gga-miR-150 and gga-miR-150mut detected by northern blotting. Positive control: spleen leukocytes from a 4-week-old chicken. Northern blots were performed with a ^{32}P -5' end-labelled DNA oligonucleotide probe complementary to the miRNA. **(B)** The MYB reporter was repressed in PA9 or MSB-1 cells producing gga-miR-150. PA9 and MSB-1 cells were electroporated with 40 μg of avian MYB WT pRL-TK reporter and 500 ng of the normalising plasmid containing the firefly luciferase gene, pcDNAMLuc. Luciferase activity was determined 24 hours after transfection. *Renilla* luciferase activity was normalised with respect to firefly luciferase activity and luciferase activity in PA9 and MSB-1 cells producing gga-miR-150mut was set at 100%. The data from one experiment representative of the three carried out are shown, with all three assays generating essentially identical results (* significant repression, $P < 0.05$, Student's *t*-test comparisons with PA9T or MSB-1T150mut).



both avian DF-1 cells and human 293FT cells. The use of target site s2 by gga-miR-150 therefore seems to result from the intrinsic properties of the chicken MYB 3'UTR. The observed functionality of avian MYB target site 2 is consistent with previous observations showing that higher levels of 3' base pairing render the site more effective [61,62]. Although both target sites, s2 and s4, display strong base pairing with the gga-miR-150 seed sequence, only site s2, which also displays base pairing for 7 nt at its 3' end, is a functional target for gga-miR-150 (Fig. 3A and 3B). This observation highlights the importance of extended 3' pairing for gga-miR-150.

Recent studies have demonstrated that MYB targeting by miR-150 has been conserved throughout evolution, as it is observed in mice, humans and zebrafish [16,37]. However, analysis of the targeting of MYB 3' UTR by miR-150 in the various species in which functional target sites have been validated has shown that miR-150 targeting seems to be dependent on the use of different target sites. Synergy between two target sites located in the last third of the 3' UTR of the human and murine *c-myb* genes is required for the regulation of these genes [16,37], whereas only one target site about 120 nt downstream from the *c-myb* stop codon is functional in chicken and zebrafish (Fig. 4). Detailed analysis of the sequence of the 3' UTR of MYB from various species showed that the short 3'UTRs of MYB (frog and

medaka) and the 3' UTR of MYB orthologs of zebrafish contained two miR-150 target sites close together in the proximal region, less than 150 nt downstream from the stop codon, and that the large 3'UTR of chickens and mammals contained two additional miR-150 target sites located close together, between 800 and 970 nt from the stop codon (Fig. 4A). However, not all these sites appeared functional and most non-functional ones were mutated, with the exception of site 4 in chicken (Fig. 4B). Site 2, for example, which is functional in chicken and zebrafish and has two conserved mutations in mammals, is not functional in mouse [37] (Fig. 4B). Surprisingly, despite the presence of sequence changes in site 2 rendering it non-functional in mammals (Fig. 4B), the 3' end sequences of this site identified in



chicken and zebrafish were found to be strongly conserved throughout evolution, from frogs to humans (Additional file 3), but were not sufficient for compensatory pairing [62]. We identified no functional role for target sites 3 and 4 in chicken MYB, whereas gga-miR-150 targeted and repressed MYB mt3 or mt4 reporter constructs but not MYB mt2 reporter construct. The mutation of target site 2 completely abolished the repressor effect of gga-miR-150, as observed with the mt1234 reporter construct, in which all target sites were mutated. By contrast, miR-150 acts at sites s3 and s4 in mice and humans (Fig. 4B)[16,37]. The mutation introduced into target site 3 of the chicken MYB, resulting in the sequence involved in base pairing to the miR-150 seed sequence being one nucleotide shorter than in mammals (Fig. 4B, Additional file 3), may account for the loss of function of this mutant site. This mutation may also affect the function of conserved site 4, which was non-functional despite a perfect seed match of 8 nt. The observed non-functionality of target site 4 suggests that this site may be functional only in synergy with target site 3, as observed in mammals [16,37] (Fig. 3 and 4).

Features common to all miR-150/MYB target sites have been identified in studies of different species: a perfect seed match of eight to nine nucleotides is required, with either additional 3' pairing or synergy with another closely located site. *c-myb* is the only member of the MYB family of transcription factor genes targeted by miR-150.

Despite the sequence conservation between avian and mammalian MYB 3'UTRs making it possible to align these sequences, our observations suggest that the intrinsic characteristics of the avian MYB 3'UTR result in avian miR-150 selecting a target site different from its mammalian orthologs.

The intermediate evolutionary position of chickens in the vertebrate group may provide insight into the evolution of MYB/miR-150 target sites. Chicken has retained the same functional target site as zebrafish and, probably, frogs. This site appears to be the principal target site in chicken, despite the presence of the putative mammalian target sites in the chicken MYB.

Conclusion

Our study of the avian MYB/gga-miR-150 interaction shows that, despite the involvement of *c-myb* in development and haematopoiesis in a wide range of organisms, from zebrafish to humans, and its regulation principally through miR-150, the conservation of miR-150 target site functionality observed between chicken and zebrafish does not extend to mammals. An additional set of functional target sites has evolved in mammals, reflecting target site adaptation in these organisms.

Methods

Cell lines

We used three chicken cell lines and one human cell line: the DF-1 chicken fibroblast cell line, the Marek's disease virus-induced lymphoma-derived MSB-1 cell line and PA9 chicken T cells, the HEK-293FT (293FT) human cell line. The DF-1 cell line was cultured in DMEM (Lonza France) supplemented with 10% foetal bovine serum and 5% chicken serum (Invitrogen-Life Technologies). The MSB-1 and PA9 cell lines were cultured in RPMI-1640 medium (Lonza France) supplemented with 10% foetal bovine serum and 5% chicken serum. The three cell lines were maintained at 41°C in an atmosphere containing 5% CO₂. The 293FT cell line (Invitrogen-Life Technologies) is a fast-growing variant of the HEK-293FT that stably expresses SV40 TAg and the neomycin resistance gene from pCMVPORT6AT. neo. 293FT cells were cultured in DMEM supplemented with 10% foetal bovine serum and were maintained at 37°C in an atmosphere containing 5% CO₂.

Plasmid construction

All primer sequences are listed in Additional file 2. The avian synthetic premiR-150 DNA sequence was obtained by overlap extension with primers 610 and 611, by polymerase chain reaction (PCR). The resulting PCR product was inserted into the pGEM-T Easy vector (Promega), digested with *KpnI* and *XhoI* and inserted into the pcDNA3 expression vector downstream from the cytomegalovirus (CMV) promoter (pmiR-150). A mutated premiR-150 DNA with a deletion of one cytosine residue from the seed sequence of miR-150 was also constructed with primers 612 and 611, using the same protocol (pmiR-150mut). The 3' untranslated region (UTR) fragment (1259 nt from the stop codon) of the avian MYB cDNA (GenBank accession number NM_205306.1), which contains four putative miR-150 target sites, was amplified by PCR with primers 682 and 683, using genomic DNA from the thymus of a four-week-old chicken as the template. The PCR product was inserted into the *NotI* site of the *Renilla* luciferase gene 3'UTR region in a pRL-TK vector (Promega), giving the MYB wild-type (WT) reporter. We generated MYB mutant reporters, containing mutations (poly-T or A-replacing sequences) generating mismatches within the "seed region" of miR-150, using a PCR-based protocol, as previously described [65]. Mutations were introduced into target site 1 directly, via the sequence of primer 806, in a one-step PCR amplification also involving primer 683. Mutations of target sites 2, 3 and 4 were introduced by overlap extension, using primer pairs 697/698, 699/700, 701/702 (Additional file 2), respectively, in an initial amplification step, followed by a second step with

primers 682 and 683 (Additional file 2). All intermediate and final constructs were checked by sequencing with appropriate primers.

Isolation of lymphoid cells ectopically producing miR-150 or miR-150mut

The stable expression of constructs encoding miR-150 or miR-150mut in PA9 cells was obtained by electroporation with the pmiR-150 and pmiR-150mut plasmids, respectively, using the Amaxa nucleofector device (Lonza): plasmid (2.5 µg) was added to 2×10^6 cells, with a pcDNA control used in parallel, and cells were cultured in complete RPMI-1640 medium without selection for 24 hours. Cells were then selected in complete RPMI-1640 medium supplemented with 1 mg/ml G418 for three weeks.

Luciferase assay

For the luciferase assay, pcDNAMLuc, carrying the firefly luciferase gene under the control of the CMV promoter, was used for cotransfection, to control for transfection efficiency.

PA-9 and MSB-1 cells were electroporated with an Equibio "Easyject Plus" electroporator (single pulse, 400 V, 1500 µF) and aluminium electrodes (4-mm cuvette, Eurogentec). For all assays, 5×10^6 cells were electroporated in the presence of 40 µg of pRL-TK_MYBWT and 500 ng of pcDNAMLuc in serum-free RPMI 1640 medium. After electroporation, cells were plated in 6-well plates, each well containing 2.5 ml of RPMI-1640 medium supplemented with serum, and cultured for 24 hours at 41°C.

DF-1 and 293FT cells were seeded in 96-well plates (2.5×10^4 cells per well). They were cultured for 24 hours before cotransfection with reporter plasmids (100 ng of pRL-TK_MYBWT or mutated and 1.25 ng of pcDNAMLuc) and 300 ng of effector plasmids (pmiR-150 or pmiR-150mut or pcDNA control), using Lipofectamine 2000 (Invitrogen-Life technologies) according to the manufacturer's protocol.

Luciferase assays were performed 24 hours after transfection, with the Dual-Luciferase Reporter assay system (Promega). *Renilla* luciferase activity was normalised with respect to firefly luciferase activity. Each transfection reaction was repeated in triplicate for each set of conditions and the experiment was carried out at least three times. Mean relative luciferase activity is presented. The significance of differences between effector and reporter constructs was assessed with Student's t-test. We considered P values < 0.05 to be statistically significant.

Northern blotting

Total RNA was extracted from 5×10^6 cells with Trizol reagent (Invitrogen-Life Technologies), according to the

manufacturer's protocol. Briefly, 15 µg of RNA was subjected to electrophoresis in a 15% acrylamide gel and transferred onto a nylon membrane (Macherey-Nagel Porablot Nylon Activated). The membrane was cross-linked by exposure to UV light (Stratalinker, Stratagene). The membrane was prehybridised for 1 h and then hybridised overnight in Perfect Hyb TM Plus hybridisation buffer (Sigma), at 50°C, with a 20 nM 32 P-5' end-labelled DNA oligonucleotide probe complementary to the miRNA, and then washed in low-stringency wash buffer. Blots were analysed by phosphorimaging with a Storm 840 (Amersham).

Sequence analyses and target prediction

The online target prediction algorithm Targetscan [66] <http://www.targetscan.org> was used to list potential miR-150 target sites from mammalian, frog and chicken MYB 3'UTRs. Additional Blast analysis of the miR-150 sequence against the MYB 3'UTR sequence of each species analysed was carried out with the accessory application "local blast" available in BioEdit version 7.0.5 sequence alignment software.

Additional material

Additional file 1: Hairpin structures of three different mammalian pre-miR-150 sequences and of the synthetic *Gallus gallus* pre-miR-150. The mature miR-150 is shown in italics and the seed sequence is shown in red.

Additional file 2: Table of primer sequences.

Additional file 3: Predicted binding structures for miR-150 and the four putative target sites (s1, s2, s3, s4) from different species, through evolution. The seed sequence (nt2 to nt8) of miR-150 is shown in red, bold. The seed sequence nucleotides involved in pairing are underlined; nucleotides involved in pairing outside the seed sequence are shown in blue, bold. The target sites nucleotides involved in pairing are shown in bold.

Acknowledgements

Audrey Guillon-Munos received postdoctoral grants from the *Ligue Nationale contre le Cancer-Comité de l'Indre*. Benoît Muylkens is a postdoctoral researcher supported by the *Fonds National de la Recherche Scientifique* (FNRS), Belgium. We thank Sébastien Pfeffer for helping to produce the small-RNA libraries and Pierre Coursaget for providing us with the 293FT cell line. This work was supported by the *Ligue Nationale contre le cancer, Comité du Cher*, and the *Agence Nationale de la Recherche* (ANR-07-MIME-012-01).

Author details

¹Université François Rabelais, Equipe Transcription, Lymphome Viro-induit, UFR Sciences et Techniques, Parc de Grandmont, F-37200 Tours, France. ²INSERM, U618 Protéases et Vectorisation Pulmonaire, UFR de Médecine, F-37000 Tours, France. ³INRA, Département de Santé Animale, F-37380 Nouzilly, France. ⁴INRA, Laboratoire de Virologie Moléculaire, UR-IASP 213, F-37380 Nouzilly, France. ⁵Université de Namur, Département Vétérinaire, Faculté des Sciences, FUNDP, 5000 Namur, Belgique.

Authors' contributions

DR conceived and designed the study. DR and GD co-ordinated the study. AGM and GD contributed to the design of the experiments. AGM, NR and DC carried out the experiments. BM screened the spleen and blood

leukocyte small-RNA libraries, isolated gga-miR-150 and helped with northern-blot analysis. AGM contributed to the interpretation of the data and to the drafting of the manuscript. GD and DR wrote the manuscript. All authors read and approved the final manuscript.

Received: 10 September 2009 Accepted: 2 September 2010
Published: 2 September 2010

References

- Klempnauer KH, Gonda TJ, Bishop JM: Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. *Cell* 1982, **31**(2 Pt 1):453-463.
- Gonda TJ, Bishop JM: Structure and transcription of the cellular homolog (c-myb) of the avian myeloblastosis virus transforming gene (v-myb). *J Virol* 1983, **46**(1):212-220.
- Lipsick JS, Baluda MA: The myb oncogene. *Gene Amplif Anal* 1986, **4**:73-98.
- Shen-Ong GL, Morse HC, Potter M, Mushinski JF: Two modes of c-myb activation in virus-induced mouse myeloid tumors. *Mol Cell Biol* 1986, **6**(2):380-392.
- Clappier E, Cucuini W, Kalota A, Crinquette A, Cayuela JM, Dik WA, Langerak AW, Montpellier B, Nadel B, Walrafen P, et al: The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. *Blood* 2007, **110**(4):1251-1261.
- Lahortiga I, De Keersmaecker K, Van Vlierberghe P, Graux C, Cauwelier B, Lambert F, Mentens N, Beverloo HB, Pieters R, Speleman F, et al: Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. *Nat Genet* 2007, **39**(5):593-595.
- Thompson MA, Flegg R, Westin EH, Ramsay RG: Microsatellite deletions in the c-myb transcriptional attenuator region associated with over-expression in colon tumour cell lines. *Oncogene* 1997, **14**(14):1715-1723.
- Hugo H, Cures A, Suraweera N, Drabsch Y, Purcell D, Mantamadiotis T, Phillips W, Dobrovic A, Zupi G, Gonda TJ, et al: Mutations in the MYB intron I regulatory sequence increase transcription in colon cancers. *Genes Chromosomes Cancer* 2006, **45**(12):1143-1154.
- Persson M, Andren Y, Mark J, Horlings HM, Persson F, Stenman G: Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. *Proc Natl Acad Sci USA* 2009, **106**(44):18740-18744.
- Guerin M, Sheng ZM, Andrieu N, Riou G: Strong association between c-myb and oestrogen-receptor expression in human breast cancer. *Oncogene* 1990, **5**(1):131-135.
- Fang F, Ryszczyn MA, Clevenger CV: Role of c-Myb during prolactin-induced signal transducer and activator of transcription 5a signaling in breast cancer cells. *Endocrinology* 2009, **150**(4):1597-1606.
- Oh IH, Reddy EP: The myb gene family in cell growth, differentiation and apoptosis. *Oncogene* 1999, **18**(19):3017-3033.
- Ramsay RG, Gonda TJ: MYB function in normal and cancer cells. *Nat Rev Cancer* 2008, **8**(7):523-534.
- Bertrand JY, Kim AD, Teng S, Traver D: CD41+ cmyb+ precursors colonize the zebrafish pronephros by a novel migration route to initiate adult hematopoiesis. *Development* 2008, **135**(10):1853-1862.
- Bertrand JY, Kim AD, Violette EP, Stachura DL, Cisson JL, Traver D: Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. *Development* 2007, **134**(23):4147-4156.
- Lin YC, Kuo MW, Yu J, Kuo HH, Lin RJ, Lo WL, Yu AL: c-Myb is an evolutionary conserved miR-150 target and miR-150/c-Myb interaction is important for embryonic development. *Mol Biol Evol* 2008, **25**(10):2189-2198.
- Nomura N, Takahashi M, Matsui M, Ishii S, Date T, Sasamoto S, Ishizaki R: Isolation of human cDNA clones of myb-related genes, A-myb and B-myb. *Nucleic Acids Res* 1988, **16**(23):11075-11089.
- Trauth K, Mutschler B, Jenkins NA, Gilbert DJ, Copeland NG, Klempnauer KH: Mouse A-myb encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes. *Embo J* 1994, **13**(24):5994-6005.
- Sala A, Watson R: B-Myb protein in cellular proliferation, transcription control, and cancer: latest developments. *J Cell Physiol* 1999, **179**(3):245-250.
- Ramsay RG: c-Myb a stem-progenitor cell regulator in multiple tissue compartments. *Growth Factors* 2005, **23**(4):253-261.
- Martin C, Paz-Ares J: MYB transcription factors in plants. *Trends Genet* 1997, **13**(2):67-73.
- Biedenkapp H, Borgmeyer U, Sippel AE, Klempnauer KH: Viral myb oncogene encodes a sequence-specific DNA-binding activity. *Nature* 1988, **335**(6193):835-837.
- Dai P, Akimaru H, Tanaka Y, Hou DX, Yasukawa T, Kanei-Ishii C, Takahashi T, Ishii S: CBP as a transcriptional coactivator of c-Myb. *Genes Dev* 1996, **10**(5):528-540.
- Oelgeschlaeger M, Janknecht R, Krieg J, Schreck S, Luscher B: Interaction of the co-activator CBP with Myb proteins: effects on Myb-specific transactivation and on the cooperativity with NF-M. *Embo J* 1996, **15**(11):2771-2780.
- Bender TP, Thompson CB, Kuehl WM: Differential expression of c-myb mRNA in murine B lymphomas by a block to transcription elongation. *Science* 1987, **237**(4821):1473-1476.
- Thomas MD, Kremer CS, Ravichandran KS, Rajewsky K, Bender TP: c-Myb is critical for B cell development and maintenance of follicular B cells. *Immunity* 2005, **23**(3):275-286.
- Garcia P, Clarke M, Vegiopoulos A, Berlanga O, Camelo A, Lorvellec M, Frampton J: Reduced c-Myb activity compromises HSCs and leads to a myeloproliferation with a novel stem cell basis. *Embo J* 2009, **28**(10):1492-1504.
- Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004, **116**(2):281-297.
- Bartel DP: MicroRNAs: target recognition and regulatory functions. *Cell* 2009, **136**(2):215-233.
- Bagga S, Bracht J, Hunter S, Massier K, Holtz J, Eachus R, Pasquinelli AE: Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* 2005, **122**(4):553-563.
- Lim LP, Lau NC, Garrett-Engle P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM: Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 2005, **433**(7027):769-773.
- Brodersen P, Voinnet O: Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol* 2009, **10**(2):141-148.
- Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M: Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* 2005, **434**(7031):338-345.
- John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS: Human MicroRNA targets. *PLoS Biol* 2004, **2**(11):e363.
- Friedman RC, Farh KK, Burge CB, Bartel DP: Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009, **19**(1):92-105.
- Zhou B, Wang S, Mayr C, Bartel DP, Lodish HF: miR-150, a microRNA expressed in mature B and T cells, blocks early B cell development when expressed prematurely. *Proc Natl Acad Sci USA* 2007, **104**(17):7080-7085.
- Xiao C, Calado DP, Galler G, Thai TH, Patterson HC, Wang J, Rajewsky N, Bender TP, Rajewsky K: MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb. *Cell* 2007, **131**(1):146-159.
- Lu J, Guo S, Ebert BL, Zhang H, Peng X, Bosco J, Pretz J, Schlanger R, Wang JY, Mak RH, et al: MicroRNA-mediated control of cell fate in megakaryocyte-erythrocyte progenitors. *Dev Cell* 2008, **14**(6):843-853.
- Barroga CF, Pham H, Kaushansky K: Thrombopoietin regulates c-Myb expression by modulating micro RNA 150 expression. *Exp Hematol* 2008, **36**(12):1585-1592.
- Tan LP, Wang M, Robertus JL, Schakel RN, Gibcus JH, Diepstra A, Harms G, Peh SC, Reijmers RM, Pals ST: miRNA profiling of B-cell subsets: specific miRNA profile for germinal center B cells with variation between centroblasts and centrocytes. *Lab Invest* 2009, **89**(6):708-716.
- Merkerova M, Belickova M, Bruchova H: Differential expression of microRNAs in hematopoietic cell lineages. *Eur J Haematol* 2008, **81**(4):304-310.
- Burnside J, Ouyang M, Anderson A, Bernberg E, Lu C, Meyers BC, Green PJ, Markis M, Isaacs G, Huang E, et al: Deep sequencing of chicken microRNAs. *BMC Genomics* 2008, **9**:185.
- Ahanda ML, Ruby T, Wittzell H, Bed'Hom B, Chausse AM, Morin V, Oudin A, Chevalier C, Young JR, Zoorob R: Non-coding RNAs revealed during identification of genes involved in chicken immune responses. *Immunogenetics* 2009, **61**(1):55-70.

44. Hicks JA, Tembhurne P, Liu HC: **MicroRNA expression in chicken embryos.** *Poult Sci* 2008, **87**(11):2335-2343.
45. Shao P, Zhou H, Xiao ZD, He JH, Huang MB, Chen YQ, Qu LH: **Identification of novel chicken microRNAs and analysis of their genomic organization.** *Gene* 2008, **418**(1-2):34-40.
46. Glazov EA, Cottee PA, Barris WC, Moore RJ, Dalrymple BP, Tizard ML: **A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach.** *Genome Res* 2008, **18**(6):957-964.
47. Rathjen T, Pais H, Sweetman D, Moulton V, Munsterberg A, Dalmay T: **High throughput sequencing of microRNAs in chicken somites.** *FEBS Lett* 2009, **583**(9):1422-1426.
48. Lambeth LS, Yao Y, Smith LP, Zhao Y, Nair V: **MicroRNAs 221 and 222 target p27Kip1 in Marek's disease virus-transformed tumour cell line MSB-1.** *J Gen Virol* 2009, **90**(Pt 5):1164-1171.
49. Zhao Y, Yao Y, Xu H, Lambeth L, Smith LP, Kgosana L, Wang X, Nair V: **A functional MicroRNA-155 ortholog encoded by the oncogenic Marek's disease virus.** *J Virol* 2009, **83**(1):489-492.
50. Shi L, Ko ML, Ko GY: **Rhythmic expression of MicroRNA-26a (mir-26a) regulates the L-type voltage-gated calcium channel [alpha]1C subunit (VGCC[alpha]1C) in chicken cone photoreceptors.** *J Biol Chem* 2009, **284**(38):25791-25803.
51. **Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution.** *Nature* 2004, **432**(7018):695-716.
52. Yao Y, Zhao Y, Xu H, Smith LP, Lawrie CH, Watson M, Nair V: **MicroRNA profile of Marek's disease virus-transformed T-cell line MSB-1: predominance of virus-encoded microRNAs.** *J Virol* 2008, **82**(8):4007-4015.
53. Yao Y, Zhao Y, Smith LP, Lawrie CH, Saunders NJ, Watson M, Nair V: **Differential expression of microRNAs in Marek's disease virus-transformed T-lymphoma cell lines.** *J Gen Virol* 2009, **90**(Pt 7):1551-1559.
54. **CSIRO-Chicken-Solexa sequencing-embryo miRNA.** [http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL6541].
55. **miRBase.** [http://microrna.sanger.ac.uk].
56. Griffiths-Jones S: **miRBase: the microRNA sequence database.** *Methods Mol Biol* 2006, **342**:129-138.
57. **Gallus gallus genome.** [http://www.ncbi.nlm.nih.gov/projects/genome/guide/chicken/].
58. Akiyama Y, Kato S: **Two cell lines from lymphomas of Marek's disease.** *Biken J* 1974, **17**(3):105-116.
59. Mazzella O, Cauchy L, Coudert F, Richard J: **Chicken thymocyte-specific antigens identified by monoclonal antibodies: characterization and distribution in normal tissues and in tumoral tissues from Marek's disease chicken.** *Hybridoma* 1986, **5**(4):319-328.
60. Chen M, Granger AJ, Vanbrocklin MW, Payne WS, Hunt H, Zhang H, Dodgson JB, Holmen SL: **Inhibition of avian leukosis virus replication by vector-based RNA interference.** *Virology* 2007, **365**(2):464-472.
61. Brennecke J, Stark A, Russell RB, Cohen SM: **Principles of microRNA-target recognition.** *PLoS Biol* 2005, **3**(3):e85.
62. Grimson A, Farh KK, Johnston WK, Garrett-Engle P, Lim LP, Bartel DP: **MicroRNA targeting specificity in mammals: determinants beyond seed pairing.** *Mol Cell* 2007, **27**(1):91-105.
63. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120**(1):15-20.
64. Himly M, Foster DN, Bottoli I, Iacovoni JS, Vogt PK: **The DF-1 chicken fibroblast cell line: transformation induced by diverse oncogenes and cell death resulting from infection by avian leukosis viruses.** *Virology* 1998, **248**(2):295-304.
65. Fragnet L, Blasco MA, Klapper W, Rasschaert D: **The RNA subunit of telomerase is encoded by Marek's disease virus.** *J Virol* 2003, **77**(10):5985-5996.
66. **Prediction of microRNA targets.** [http://www.targetscan.org].

doi:10.1186/1471-2199-11-67

Cite this article as: Guillon-Munos *et al.*: The chicken miR-150 targets the avian orthologue of the functional zebrafish MYB 3'UTR target site. *BMC Molecular Biology* 2010 **11**:67.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit





Damien COUPEAU

Etude transcriptionnelle et post-transcriptionnelle de la région IRL/TRL du virus de la maladie de Marek

Résumé

Le Gallid Herpesvirus 2 (GaHV-2) est un herpesvirus responsable de lymphomes T chez le poulet. Son génome code pour 25 miARN matures regroupés en 2 clusters localisés au sein des régions répétées TR_L/IR_L et IR_S/TR_S. Mon travail a tout d'abord consisté en l'étude de la fonctionnalité des deux brins de mdv1-pré-miR-M4 sur deux cibles virales UL-28 et UL32. Le brin mdv1-miR-M4-5P ayant été identifié comme orthologue du miARN cellulaire miR-155, leurs impacts ont été comparés sur six cibles cellulaires et la cible virale UL-28. Les différents miARN de GaHV-2 présentant des taux d'expression variables, nous avons identifié, dans la région TR_L/IR_L, des ARNm responsables de leur transcription à partir d'exon ou d'intron. De plus, nous avons identifié un promoteur plutôt spécifique de la latence et responsable de la transcription de l'ensemble des miARN de l'IR_L/TR_L, de l'oncogène *meq* et du transcrit Meq/vIL-8.

Mots clefs : Herpesvirus, GaHV-2, TR_L/IR_L, microARN, épissage, promoteur, édition

Abstract

Gallid Herpesvirus 2 (GaHV-2) is a herpesvirus inducing T-lymphoma in chickens. Its genome encodes 25 mature miRNAs splitted in two clusters localized in IR_L/TR_L and IR_S/TR_S repeated regions. In this work, we first analyzed the functionality of the two strands of mdv1-pre-miR-M4 on two viral targets named UL-28 and UL-32. As one strand of this miRNA was identified as an ortholog of cellular miRNA-155, their effects were assessed on six cellular targets and on the UL-28 viral target. As a differential expression of miRNAs was observed during GaHV-2 infection, we found numerous IR_L/TR_L transcripts responsible for the intronic or exonic transcription of viral miRNAs. Moreover, we identified a latency promoter controlling at the same time transcription of all IR_L/TR_L miRNAs, the oncogenic *meq* gene and the Meq/v-IL8 transcript.

Key words : Herpesvirus, GaHV-2, IR_L/TR_L, microRNA, splicing, promoter, editing.